

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 November 2001 (08.11.2001)

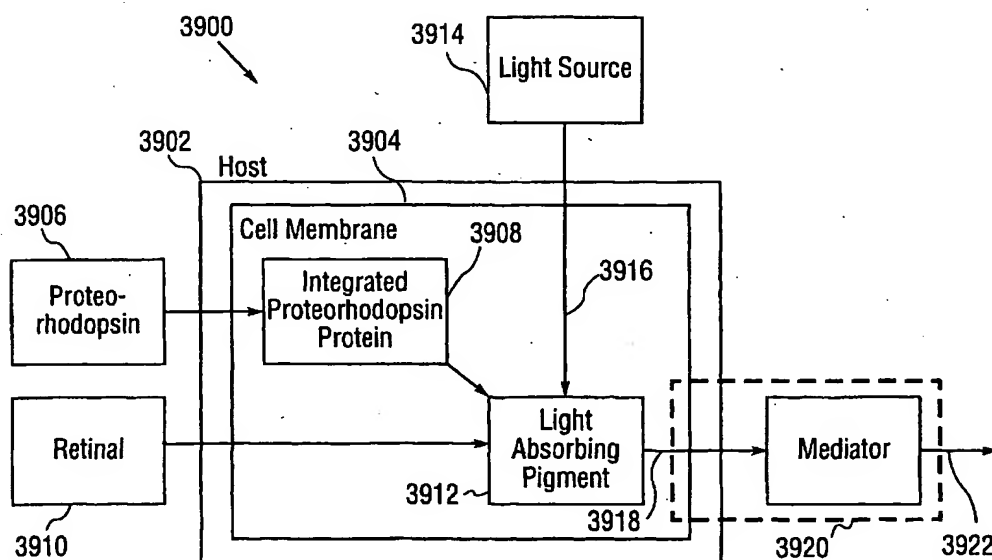
PCT

(10) International Publication Number
WO 01/83701 A2

- (51) International Patent Classification⁷: C12N (74) Agent: KAMINSKY, Rena; Lumen Intellectual Property Services, Inc., 45 Cabot Avenue, Suite 110, Santa Clara, CA 95051 (US).
- (21) International Application Number: PCT/US01/14394
- (22) International Filing Date: 2 May 2001 (02.05.2001) (81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CO, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW.
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/201,602 3 May 2000 (03.05.2000) US (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant: MONTEREY BAY AQUARIUM RESEARCH INSTITUTE [US/US]; 7700 Sandholt Road, Moss Landing, CA 95039 (US).
- (72) Inventors: DELONG, Edward, F.; 4 Blacktail Lane, Monterey, CA 93940 (US). BEJA, Oded; 408 Holovits Court, Marina, CA 93933 (US). Published: — without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: LIGHT-DRIVEN ENERGY GENERATION USING PROTEORHODOPSIN



(57) Abstract: A light-driven energy generation system using proteorhodopsin is provided. Proteorhodopsin sequences were retrieved and amplified from naturally occurring members of the domain Bacteria using proteorhodopsin-specific polymerase chain reaction primers. Proteorhodopsin sequences were placed in expression vectors for production of proteorhodopsin proteins in a host, for instance, *E. coli* and other bacteria. The system also includes a light source and a source of retinal, that allows the system to convert light into biochemical energy. The generated biochemical energy could be mediated into electrical energy by a mediator.

BEST AVAILABLE COPY



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



PATENT APPLICATION

LIGHT-DRIVEN ENERGY GENERATION USING PROTEORHODOPSIN

INVENTORS

Edward F. DeLong and Oded Beja

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is cross-referenced to and claims priority from U.S. Provisional application' 60/201,602 filed 05/03/2000, which is hereby incorporated by reference.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention was supported in part by grant number OCE 0001619 from the National Science Foundation (NSF). The U.S. government has certain rights in the invention.

STATEMENT TO COMPUTER DISK AND SEQUENCE LISTING

This application includes a sequence listing of 65 sequences and a computer disk labeled "Sequence Listing for application entitled "Light-driven energy generation using proteorhodopsin" by Edward F. DeLong and Oded Beja" containing files "MBA101-SEQLIST.prj", dated "04/23/01" with 174,089 bytes, which is the PatentIn

project file generated using PatentIn Version 3.0 software provided by the USPTO, and "MBA101-SEQLIST.txt", dated "04/23/01" with 323,739 bytes, which is the generated sequence listing from the PatentIn project file MBA101-SEQLIST.prj using PatentIn Version 3.0 software, all which are herein incorporated. The information recorded in computer readable format on the incorporated computer disk labeled "Sequence Listing" containing files "MBA101-SEQLIST.prj" and "MBA101-SEQLIST.txt" are identical to the incorporated written sequence listing.

FIELD OF THE INVENTION

The present invention relates generally to gene expression of functional recombinant proteins in bacteria. More particularly, the present invention relates to proteorhodopsin genes and proteins that function as a light-driven energy generator in *Escherichia coli* (E. coli) and other bacteria.

BACKGROUND ART

Retinal (vitamin A aldehyde) is a chromophore that binds integral membrane proteins (opsins) to form light-absorbing pigments called rhodopsins. Rhodopsins are currently known to belong to two distinct protein families. The visual rhodopsins, found in the eye throughout the animal kingdom, are photosensory pigments. Archeal rhodopsins, found in extreme halophilic environments, function as light-driven protons pumps (bacteriorhodopsins), chloride ion pumps (halorhodopsins), or photosensory receptors (sensory rhodopsins). The two protein families show no significant sequence similarity and may have different origins. They do, however, share identical topologies characterized by seven transmembrane α -helices that form a pocket in which retinal is covalently linked, as a protonated Schiff base (helix G).

The archaeal rhodopsins are able to generate a photocycle which produces a chemiosmotic membrane potential in response to light, as such light energy is converted into biochemical energy. Recently, a protein with high sequence similarity to the archaeal rhodopsins has also been retrieved in the eukaryote *Neurospora crassa* (J.A. Bieszke et al., *Proceedings of National Academy of Sciences USA* 96:8034, 1999). The eucaryal rhodopsin formed a photochemically reactive pigment when bound to all-trans retinal and exhibited photocycle kinetics similar to those of archaeal sensory rhodopsins (J.A. Bieszke et al., *Biochemistry* 38:14138, 1999). To date, however, no rhodopsin-like sequences have been reported in members of the domain Bacteria, and no light-driven proton pumps based on rhodopsin have ever before been functionally expressed in *E. coli*.

The phototropic conversion of light energy into biochemical energy using bacteriorhodopsin can be harnessed for a variety of processes and applications, such as bio-electronic applications and bio-materials, as has been reported in US Patent No. 5,757,525 for optical devices, US Patent No. 5,854,710 for optical Fourier processing, and US Patent No. 5,470,690 for optical information storage. Bacteriorhodopsin in bio-electronic applications is aimed to interface, integrate, or substitute the silicon based microelectronics systems as well as molecular devices. Bacteriorhodopsin as a bio-material is integrated, for instance, in optical films for light mediated computer memory applications and pattern recognition.

Previously, archaeal rhodopsins capable of generating a chemiosmotic membrane potential in response to light had only been found in halophilic archaea. Therefore, rhodopsins that originate from archaea adapted to highly saline environments cannot be functionally expressed in *E. coli*. Finally, the isolation and cultivation of

halorhodopsins is an elaborate process. At present one does not foresee an economic utilization possible for this process (e.g. US Patent 5,290,699).

Accordingly, as one skilled in the art might readily acknowledge, there is a strong need to retrieve and provide rhodopsin-like sequences from naturally occurring members of the domain Bacteria.

OBJECTS AND ADVANTAGES

In light of the above, it is the primary objective of the present invention to provide rhodopsin-like sequences from naturally occurring members of the domain Bacteria. More specifically, it is the objective of the present invention to provide a method to retrieve proteorhodopsin genes from DNA of naturally occurring bacteria that encodes DNA sequence for proteorhodopsin proteins.

It is another objective of the present invention to provide proteorhodopsin-specific polymerase chain reaction primers that amplify the proteorhodopsin-containing gene from a DNA sample of naturally occurring bacteria.

It is yet another objective of the present invention to produce variants of a proteorhodopsin gene using the same proteorhodopsin-specific polymerase chain reaction primers by amplifying a proteorhodopsin-containing gene from of a mixed sample of naturally occurring bacteria.

It is still another objective of the present invention to provide an expression vector that produces a proteorhodopsin protein in *E. coli* and other bacteria.

It is another objective of the present invention to provide a light-driven energy generator in which the functional properties of proteorhodopsin are utilized. These properties include the ability to integrate within a host, for instance a cell membrane of *E. coli*, making an integrated proteorhodopsin protein, and the ability to bind retinal, making a light absorbing pigment.

It is another objective of the present invention to provide a light source and illuminate the light absorbing pigment to convert light energy into biochemical energy.

It is another objective of the present invention to provide a mediator and mediate the biochemical energy into electrical energy.

It is another objective of the present invention to provide methods to manipulate the kinetics of the light-driven energy generator.

The advantage of the present invention over the prior art is that it is not restricted to operate in halophilic archaea and could therefore be functionally expressed in *E. coli* and other bacteria. Accordingly, another advantage of the present invention is that it provides for a fast and cheap production method that allows for mass production of functionally active proteorhodopsin.

SUMMARY

The present invention provides proteorhodopsin gene and protein sequences retrieved from samples of naturally occurring members of the domain Bacteria. More specifically, the present invention provides a method for the retrieval and amplification of proteorhodopsin genes from DNA samples of naturally occurring marine bacteria. In accordance with several exemplary embodiments of the present invention, DNA samples were obtained from naturally occurring bacteria such as, for instance, marine proteobacteria, SAR86 bacteria, or recombinant DNA libraries containing naturally occurring bacteria. The present invention provides proteorhodopsin-specific polymerase chain reaction (PCR) primers to amplify a proteorhodopsin gene from DNA samples of these marine bacteria. The present invention also provides a device and method for the placement of proteorhodopsin genes in an expression vector to produce functional proteorhodopsin proteins in *E. coli* and other bacteria.

Accordingly, the present invention provides a method to produce and obtain variants of proteorhodopsin genes and proteins. The same proteorhodopsin-specific polymerase chain reaction primers amplify different variants of proteorhodopsin-containing genes from a mixed sample of naturally occurring bacteria. As one skilled in the art might readily acknowledge, these variants of a proteorhodopsin gene produce functional variations in the photocycle kinetics of the proteorhodopsin protein.

Furthermore, the present invention provides a light-driven energy generator that utilizes proteorhodopsin to convert light-energy into biochemical energy. This light-driven energy generator takes advantage of the functional properties of the proteorhodopsin protein once expressed in, for example, *E. coli* or other bacteria as is

described in exemplary embodiments. These properties include the ability to integrate within a host such as, for instance, a cell membrane of *E. coli* or other Bacteria, and thereby making an integrated proteorhodopsin protein or integrated cell membrane protein. These properties also include the ability to bind retinal and thereby making a light absorbing pigment. Illuminating the light absorbing pigment with a light source converts light energy into biochemical energy. Finally, the biochemical energy can be mediated into electrical energy by a mediator.

In accordance with exemplary embodiments, the present invention enables one skilled in the art to manipulate the kinetics of the proteorhodopsin protein photocycle once it is operational in the light-driven energy generator. In particular, the present invention provides examples in which the light source characteristics are manipulated. Examples are the manipulation of the delivery of fast-light pulses and/or the delivery of light at different wavelengths. The present invention also provides examples in which incremental additions of retinal influences the function of the light-driven energy generator. In addition, a proteorhodopsin gene or protein variant can be selected to determine an absorption spectra of the light absorbing pigment to change the kinetics of the light energy generator, for instance to meet a design/functional criteria of an application wherein proteorhodopsin is utilized.

BRIEF DESCRIPTION OF THE FIGURES

The objectives and advantages of the present invention will be understood by reading the following detailed description in conjunction with the drawings, in which:

- FIG. 1** illustrates the phylogenetic tree of bacterial 16S rRNA gene sequences including that encoded on the 130 kb bacterioplankton BAC clone (EBAC31A8).
- FIG. 2** provides a nucleotide sequence of polymerase chain reaction primer 1 (Sequence ID No:2) used to amplify a proteorhodopsin gene.
- FIG. 3** provides a nucleotide sequence of polymerase chain reaction primer 2 (Sequence ID No:3) used to amplify a proteorhodopsin gene.
- FIG. 4** provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:4) amplified from clone EBAC31A8 (Sequence ID No:1) using PCR primers 1 (Sequence ID No:2) and 2 (Sequence ID No:3), and the deduced amino acid sequence (Sequence ID No:5) of the proteorhodopsin gene Sequence ID No:4 amplified from clone EBAC31A8 (Sequence ID No:1).
- FIG. 5** provides a map of the secondary structure of the proteorhodopsin protein (Sequence ID No:7). Single letter amino acid codes are used (according to J. Sasaki and J.L. Spudich, Biophys. J. 75:2435, 1998). Predicted retinal binding pocket residues are marked in black.
- FIG. 6** provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:8) amplified from clone EBAC40E8 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:9) of the proteorhodopsin gene Sequence ID No:8 amplified from clone EBAC40E8.
- FIG. 7** provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:10) amplified from clone EBAC41B4 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:11) of the proteorhodopsin gene Sequence ID No:7 amplified from clone EBAC41B4.

FIG. 8 provides the nucleotide sequence of the proteorhodopsin gene (**Sequence ID No:12**) amplified from clone EBAC64A5 using PCR primers 1 (**Sequence ID No:2**) and 2 (**Sequence No:3**), and the deduced amino acid sequence (**Sequence ID No:13**) of the proteorhodopsin gene **Sequence ID No:12** amplified from clone EBAC64A5.

FIG. 9 provides a variants map of the DNA sequences of the proteorhodopsin gene with **Sequence ID No:4**, **Sequence ID No:8**, **Sequence ID No:10**, and **Sequence ID No:12** that were amplified from clone EBAC38A8, EBAC40E8, EBAC41B4 and EBAC64A5 respectively using the proteorhodopsin-specific PCR primer 1 (**Sequence ID No:2**) and 2 (**Sequence ID No:3**). Dots represent sequences having identical sequence as those in **Sequence ID No:4**.

FIG. 10 provides a variant map of the deduced amino acid sequences encoded by the proteorhodopsin gene with **Sequence ID No:4**, **Sequence ID No:8**, **Sequence ID No:10**, and **Sequence ID No:12** that were amplified from respectively EBAC38A8, EBAC40E8, EBAC41B4 and EBAC64A5 using the proteorhodopsin-specific primer 1 (**Sequence ID No:2**) and 2 (**Sequence ID No:3**). Lower case represents the PCR primer sequence region. Dots represent residues having identical sequence as those in **Sequence ID No:5**.

FIG. 11 provides the nucleotide sequence of the proteorhodopsin gene (**Sequence ID No:14**) amplified from clone HOT0m1 using PCR primers 1 (**Sequence ID No:2**) and 2 (**Sequence No:3**), and the deduced amino acid sequence (**Sequence ID No:15**) of the proteorhodopsin gene **Sequence ID No:14** amplified from clone HOT0m1.

FIG. 12 provides the nucleotide sequence of the proteorhodopsin gene (**Sequence ID No:16**) amplified from clone HOT75m1 using PCR primers 1 (**Sequence ID**

No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:17) of the proteorhodopsin gene Sequence ID No:16 amplified from clone HOT75m1.

FIG. 13 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:18) amplified from clone HOT75m3 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:19) of the proteorhodopsin gene Sequence ID No:18 amplified from clone HOT75m3.

FIG. 14 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:20) amplified from clone HOT75m4 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:21) of the proteorhodopsin gene Sequence ID No:20 amplified from clone HOT75m4.

FIG. 15 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:22) amplified from clone HOT75m8 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:23) of the proteorhodopsin gene Sequence ID No:22 amplified from clone HOT75m8.

FIG. 16 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:24) amplified from clone MB0m1 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:25) of the proteorhodopsin gene Sequence ID No:24 amplified from clone MB0m1.

FIG. 17 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:26) amplified from clone MB0m2 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence

(Sequence ID No:27) of the proteorhodopsin gene Sequence ID No:26 amplified from clone MB0m2.

FIG. 18 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:28) amplified from clone MB20m2 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:29) of the proteorhodopsin gene Sequence ID No:28 amplified from clone MB20m2.

FIG. 19 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:30) amplified from clone MB20m5 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:31) of the proteorhodopsin gene Sequence ID No:30 amplified from clone MB20m5.

FIG. 20 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:32) amplified from clone MB20m12 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:33) of the proteorhodopsin gene Sequence ID No:32 amplified from clone MB20m12.

FIG. 21 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:34) amplified from clone MB40m1 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:35) of the proteorhodopsin gene Sequence ID No:34 amplified from clone MB40m1.

FIG. 22 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:36) amplified from clone MB40m5 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence

(Sequence ID No:37) of the proteorhodopsin gene Sequence ID No:36 amplified from clone MB40m5.

FIG. 23 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:38) amplified from clone MB40m12 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:39) of the proteorhodopsin gene Sequence ID No:38 amplified from clone MB40m12.

FIG. 24 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:40) amplified from clone MB100m5 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:41) of the proteorhodopsin gene Sequence ID No:40 amplified from clone MB100m5.

FIG. 25 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:42) amplified from clone MB100m7 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:43) of the proteorhodopsin gene Sequence ID No:42 amplified from clone MB100m7.

FIG. 26 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:44) amplified from clone MB100m9 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:45) of the proteorhodopsin gene Sequence ID No:44 amplified from clone MB100m9.

FIG. 27 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:46) amplified from clone MB100m10 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence

(Sequence ID No:47) of the proteorhodopsin gene Sequence ID No:46 amplified from clone MB100m10.

FIG. 28 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:48) amplified from clone PALB1 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:49) of the proteorhodopsin gene Sequence ID No:48 amplified from clone PALB1.

FIG. 29 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:50) amplified from clone PALB2 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:51) of the proteorhodopsin gene Sequence ID No:50 amplified from clone PALB2.

FIG. 30 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:52) amplified from clone PALB5 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:53) of the proteorhodopsin gene Sequence ID No:52 amplified from clone PALB5.

FIG. 31 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:54) amplified from clone PALB7 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:55) of the proteorhodopsin gene Sequence ID No:54 amplified from clone PALB7.

FIG. 32 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:56) amplified from clone PALB6 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence

(Sequence ID No:57) of the proteorhodopsin gene Sequence ID No:56 amplified from clone PALB6.

FIG. 33 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:58) amplified from clone PALB8 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:59) of the proteorhodopsin gene Sequence ID No:58 amplified from clone PALB8.

FIG. 34 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:60) amplified from clone PALE1 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:61) of the proteorhodopsin gene Sequence ID No:60 amplified from clone PALE1.

FIG. 35 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:62) amplified from clone PALE6 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:63) of the proteorhodopsin gene Sequence ID No:62 amplified from clone PALE6.

FIG. 36 provides the nucleotide sequence of the proteorhodopsin gene (Sequence ID No:64) amplified from clone PALE7 using PCR primers 1 (Sequence ID No:2) and 2 (Sequence No:3), and the deduced amino acid sequence (Sequence ID No:65) of the proteorhodopsin gene Sequence ID No:64 amplified from PALE7.

FIG. 37 illustrates a phylogenetic tree of different proteorhodopsin genes.

FIG. 38 provides an example of an alignment of proteorhodopsin amino acid sequences.

FIG. 39 provides a light-driven energy generator that utilizes proteorhodopsin.

- FIG. 40 provides an example of a proteorhodopsin-expressing *E. coli* cell suspension (+) compared to control cells (-), both with all-*trans* retinal.
- FIG. 41 provides an example of absorption spectra of retinal-constituted proteorhodopsin in *E. coli* membranes and a negative control.
- FIG. 42 provides an example of a light-driven transport of protons by a proteorhodopsin-expressing *E. coli* cell suspension.
- FIG. 43 provides an example of a transport of [³H]TPP⁺ in *E. coli* right-side-out vesicles containing expressed proteorhodopsin, reconstituted with or without 10 μM retinal in the presence of light or in the dark.
- FIG. 44 provides an example of laser flash-induced absorbance changes in suspensions of *E. coli* membranes containing proteorhodopsin.
- FIG. 45 provides an example of absorption spectra of retinal-constituted proteorhodopsin in *E. coli* membranes.

DETAILED DESCRIPTION

Although the following detailed description contains many specifics for the purposes of illustration, anyone of ordinary skill in the art will appreciate that many variations and alterations to the following details are within the scope of the invention. Accordingly, the following preferred embodiment of the invention is set forth without any loss of generality to, and without imposing limitations upon, the claimed invention.

Proteorhodopsin

The present invention provides rhodopsin-like gene and protein sequences retrieved from naturally occurring members of the domain Bacteria. More specifically, the

present invention provides a method for the retrieval and amplification of proteorhodopsin genes from DNA samples of naturally occurring marine bacteria. In accordance with exemplary embodiments of the present invention, DNA samples were obtained from naturally occurring marine bacteria such as bacteria from the SAR86 group. Provided as an exemplary embodiment of the SAR86 group, DNA samples were obtained from a bacterioplankton Bacterial Artificial Chromosome (BAC) clone BAC31A8 (also referred to as EBAC31A08). In general, as will be appreciated by those of ordinary skill in the art, suitable DNA samples can also be obtained from other sources, e.g., from a marine environment or from a recombinant DNA library containing genomic fragments of samples of naturally occurring bacteria.

FIG. 1 shows the phylogenetic tree of bacterial 16S rRNA gene sequences including that encoded on the EBAC31A8. **FIG. 1** also shows the relationship of EBAC31A8 to the SAR86 bacteria group as well as to the gamma-proteobacteria group. A subclone shotgun library was constructed from BAC clone 31A8, and subclones were sequenced in both directions on the MegaBACE 1000 capillary array electrophoresis DNA sequencing instrument (Molecular Dynamics, Sunnyvale, CA). Sequence analysis of a 130-kb genomic DNA that encodes the ribosomal RNA operon from BAC31A8, reveals an open reading frame encoding a proteorhodopsin. In an exemplary embodiment, the contiguous sequence was assembled using SEQUENCHER 3.1.1 software (Gene Codes Co., Ann Arbor, MI). Other sequencing techniques can also be used, as will be recognized by those skilled in the art. The sequence of the proteorhodopsin-containing contig has been deposited in GenBank under accession #AF279106 and deposit date October 23rd, 2000. **Appendix A**, hereby incorporated, shows the nucleotide sequence of the BAC clone BAC31A8 (**Sequence ID No:1**)

which contains the 130 kilobases genomic DNA from a naturally occurring marine bacterium.

Proteorhodopsin was amplified from the 130 kilobase bacterioplankton BAC clone 31A8 (**Sequence ID No:1**) by polymerase chain reaction (PCR), using the proteorhodopsin-specific primers 5'-aCCATGGgtaaattattactgatattagg-3' (**Sequence ID No:2** and shown in **FIG. 2**) and 5'-agcattagaagattctttaacagc-3' (**Sequence ID No:3** and shown in **FIG. 3**). References for PCR are, for instance, *The Polymerase Chain Reaction*, Mullis et al., Ed. (Birkhauser, Boston, 1994) and U.S. Patent Nos. 4,683,195 and 4,683,202 to Mullis et al. The proteorhodopsin-specific PCR primers include the addition of 3 nucleotides that encoded one amino acid not found in the native gene sequence of clone BAC31A8 (**Sequence ID No:6**), in the second amino acid position which is a glycine located on the 2nd codon ("GGT"). Therefore, compare the second amino acid position in the **Sequence ID No:5** using PCR primers 1 and 2 with the native **Sequence ID no:7**. This addition of one non-native amino acid created a new restriction endonuclease site (NcoI site) not present in the native sequence. This allowed subcloning of the amplified fragment into the NcoI restriction site of an expression vector pBAD TOPO TA Cloning[®] Kit (Invitrogen, La Jolla, CA). The present invention is not limited to the use of this type of expression vector and other expression vectors could also be used.

FIG. 4 shows the nucleotide sequence of the proteorhodopsin gene (**Sequence ID No:4**) that results from amplification of the proteorhodopsin-containing DNA in BAC31A8 using proteorhodopsin-specific PCR primers **Sequence ID No:2** and **Sequence No:3**. **FIG. 4** also shows the deduced amino acid sequences (**Sequence ID No:5**) encoded by the proteorhodopsin gene (**Sequence ID No:4**).

FIG. 5 shows an exemplary embodiment of a secondary structure of proteorhodopsin after it has been folded in a cell membrane **510** and bonded with retinal **520**. **FIG. 5** shows the native proteorhodopsin gene (**Sequence ID No:6**) obtained from clone BAC31A8 and encodes a proteorhodopsin protein of 249 amino acids with a molecular weight of 27 kD (**Sequence ID No:7**). In **FIG. 5**, **530** indicates seven transmembrane domains, a typical feature of the rhodopsin protein family, that aligned well with the corresponding helices of the archaeal rhodopsins. **FIG. 5** also shows the amino acid residues that form a retinal binding pocket indicated by **520**. Although the proteorhodopsin proteins shown in **FIGS. 4** and **5** both originate from BAC31A8, they differ with respect to the second amino acid position. The reason is that the proteorhodopsin-specific PCR primers that were used to amplify the proteorhodopsin gene from BAC31A8 (which resulted in proteorhodopsin protein as in **FIG. 4**; **Sequence ID No:5**) included the addition of 3 nucleotides. These 3 nucleotides encoded one amino acid not found in the native gene sequence (**Sequence ID No:6**), in the second amino acid position which is a glycine located on the 2nd codon ("GGT"). Proteorhodopsin protein (**Sequence ID No:7**) as shown in **FIG. 5** originates from the native gene sequence without the addition of the 3 nucleotides. As mentioned above, the addition of the 3 nucleotides created a new restriction endonuclease site (NcoI site) that was not present in the native sequence and thereby allowed the amplified fragment to be subcloned into the NcoI site of the expression vector.

In the exemplary embodiment presented above, PCR primers with **Sequence ID No:2** and **Sequence ID No:3** were used. In general, the present invention provides a method for designing different proteorhodopsin-specific PCR primers that are all capable of amplifying a proteorhodopsin gene from DNA samples of naturally occurring microbial populations by polymerase chain reaction. In designing these

primers one first needs to determine a DNA sequence of a proteorhodopsin gene. Then one can design oligodeoxynucleotide primers with a Watson-Crick base pair complementary to 5' and 3' ends of the proteorhodopsin gene.

Variants of Proteorhodopsin

In the previous section, an exemplary embodiment is provided of a proteorhodopsin gene and protein. The present invention also provides the retrieval of genetic variations of proteorhodopsin from naturally occurring genetic variations in naturally occurring bacterial populations. These genetic variations in proteorhodopsin sequences result in functional variations in the proteorhodopsin proteins as is discussed below.

The present invention enables one skilled in the art to use the same proteorhodopsin-specific PCR primers as shown in **FIGS. 2 and 3** to successfully amplify different sequence variants from DNA originating from mixed naturally occurring bacterial populations when it is compared to for instance the proteorhodopsin gene as shown in **FIG. 4**. As mentioned above, different proteorhodopsin-specific PCR primers could be used to amplify genetic variants of proteorhodopsin.

FIGS. 6-8 show exemplary embodiments of three different and unique variants of the proteorhodopsin gene that were retrieved from a recombinant DNA library of other naturally occurring bacteria (i.e. the bacterial artificial chromosome library (BAC)). In general, genetic variants could be obtained from different DNA libraries containing naturally occurring bacteria as well as from samples of naturally occurring bacteria. **FIG. 6** shows the variant of the proteorhodopsin gene sequence (**Sequence ID No:8**) that is amplified from the BAC clone 40 (BAC40E8) with the same proteorhodopsin-

specific PCR primers as provided in **Sequence ID No:2** and **3**. Accordingly, **FIG. 6** also shows the deduced amino acid sequence (**Sequence ID No:9**) of the genetic variant of proteorhodopsin shown in **FIG. 6**. **FIG. 7** shows the variant of the proteorhodopsin gene sequence (**Sequence ID No:10**) that is amplified from the BAC clone 41 (BAC41B4) with the same proteorhodopsin-specific PCR primers as provided in **Sequence ID No:2** and **3**. Accordingly, **FIG. 7** also shows the deduced amino acid sequence (**Sequence ID No:11**) of the genetic variant of proteorhodopsin shown in **FIG. 7**. **FIG. 8** shows the variant of the proteorhodopsin gene sequence (**Sequence ID No:12**) that is amplified from the BAC clone 64 (BAC64A5) with the same proteorhodopsin-specific PCR primers as provided in **Sequence ID No:2** and **3**. Accordingly, **FIG. 8** also shows the deduced amino acid sequence (**Sequence ID No:13**) of the genetic variant of proteorhodopsin shown in **FIG. 8**.

FIG. 9 provides a variants map of the nucleotide sequences of the proteorhodopsin gene **Sequence ID No:4**, **Sequence ID No:8**, **Sequence ID No:10**, and **Sequence ID No:12** amplified from respectively BAC31A8, BAC40E8, BAC41B4 and BAC64A5 using the proteorhodopsin-specific PCR primers **Sequence ID No:2** and **Sequence ID No:3**. In **FIG. 9** lower case letters represent the PCR primer sequence region. Dots represent residues having identical sequence as those in **Sequence ID No:4**. These proteorhodopsin gene sequences differ by as much as 31 nucleotides as is shown in **FIG. 10**. **FIG. 10** provides a variant map of the deduced amino acid sequences of the proteorhodopsin genes shown in **FIG. 9**.

Using the same proteorhodopsin-specific PCR primers, as for instance shown in **FIGS. 2** and **3**, proteorhodopsin genes were also amplified from bacterioplankton extracts. As mentioned above, any proteorhodopsin-specific PCR primer can be used. These bacterioplankton extracts include those from the Monterey Bay (referred to as MB

clones), the Southern Ocean (Palmer Station, referred to as PAL clones), and waters of the central North Pacific Ocean (Hawaii Ocean Time series station, referred to as HOT clones).

FIGS. 11-36 show exemplary embodiments of different and unique variants of proteorhodopsin that were retrieved from the MB clones, PAL clones, and HOT clones. **FIGS. 11-36** each show a variant of a proteorhodopsin gene sequence that is amplified with the same proteorhodopsin-specific PCR primers as provided in **Sequence ID No:2** and **Sequence ID No:3** from respectively clones HOT0m1, HOT75m1, HOT75m3, HOT75m4, HOT75m8, MB0m1, MB0m2, MB20m2, MB20m5, MB20m12, MB40m1, MB40m5, MB40m12, MB100m5, MB100m7, MB100m9, MB100m10, PALB1, PALB2, PALB5, PALB7, PALB6, PALB8, PALE1, PALE6 and PALE7. The proteorhodopsin gene sequences retrieved from clones HOT0m1, HOT75m1, HOT75m3, HOT75m4, HOT75m8, MB0m1, MB0m2, MB20m2, MB20m5, MB20m12, MB40m1, MB40m5, MB40m12, MB100m5, MB100m7, MB100m9, MB100m10, PALB1, PALB2, PALB5, PALB7, PALB6, PALB8, PALE1, PALE6 and PALE7, have respectively **Sequence ID Nos: 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, and 64**. Accordingly, **FIGS. 11-36** also show the deduced amino acid sequence of each genetic variant of proteorhodopsin. The deduced amino acid sequence encoded by the proteorhodopsin gene retrieved from clones HOT0m1, HOT75m1, HOT75m3, HOT75m4, HOT75m8, MB0m1, MB0m2, MB20m2, MB20m5, MB20m12, MB40m1, MB40m5, MB40m12, MB100m5, MB100m7, MB100m9, MB100m10, PALB1, PALB2, PALB5, PALB7, PALB6, PALB8, PALE1, PALE6 and PALE7, have respectively **Sequence ID Nos: 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, and 65**.

In an exemplary embodiment shown in **FIG. 37**, fifteen different variants of proteorhodopsin in the PCR generated MB gene library **3710** were detected, falling into three clusters. The MB gene library includes MB clones MB0m2, MB40m5, MB20m2, MB40m12, MB100m10, MB20m12, MB40m1, MB100m5, MB20m5, MB100m7, MB0m1, and MB100m9 as well as BAC clones BAC40E8, BAC31A8 and BAC64A5. **FIG. 37** is based on a phylogenetic analysis of the inferred amino acids of cloned proteorhodopsin genes. Evolutionary distances calculated from 220 positions were used to infer the tree topology by the neighbor joining method using the PaupSearch program of the Wisconsin Package version 10.0 (Genetics Computer Group (GCG), Madison Wisconsin). Other methods could also be used. The variants of the MB library share at least 97% identity over 248 amino acids, as shown in **FIG. 38**, and 93% identity at the DNA level. All the PCR amplified proteorhodopsin genes from Antarctic marine bacterioplankton (e.g. the PAL clones) were different from those of Monterey Bay (e.g. the MB clones) sharing 78% identity over 248 amino acids with the Monterey clade. The changes in amino acid sequences were not restricted to the hydrophilic loops, but spread over the entire protein including changes near the retinal binding domain **3830** as shown in **FIG. 38**, which are predicted retinal-binding residues. **FIG. 38** shows an example of a multiple alignment of proteorhodopsin amino acid sequences that were obtained from different clones **3820**. The secondary structure is derived from hydropathy plots (boxes **3810** shows trans-membrane helices).

Light-driven energy generator

FIG. 39 provides a light-driven energy generator **3900** that utilizes proteorhodopsin, as obtained from naturally occurring bacteria as described above, to convert light-energy into biochemical energy. Light-driven energy generator **3900** takes advantage of the

functional properties of the proteorhodopsin protein once expressed in, for instance, *E. coli* and other bacteria. These properties include the ability of proteorhodopsin 3906 to integrate within the cell membrane 3904 of, for instance, *E. coli* making an integrated proteorhodopsin protein 3908 (also called an integrated cell membrane protein). These properties also include the ability of proteorhodopsin 3906 to bind retinal 3910, making a light absorbing pigment 3912. The source of retinal 3910 is not limited to chromophore retinal but could also include chemical derivatives of retinal, such as 3-methyl-5-(1-pyryl)-2E,4E-pentadienal, 3,7-dimethyl-9-(1-pyryl)-2E,4E,6E,8E-nonatetraenal, all-trans-9-(4-azido-2,3,5,6-tetrafluorophenyl)-3,7-dimethyl-2,4,6,8,-nonatetraenal and 2,3-dehydro-4-oxoretinal. Illuminating light absorbing pigment 3912 with a light source 3914 results in a chemiosmotic gradient or proton pump in which light energy 3916 is converted into biochemical energy 3918. The chemiosmotic gradient involves pumping of protons from the inside to the outside of cell membrane 3904. When the protons return to the inside of cell membrane 3904 it produces biochemical energy 3918 via a proton translocating ATP-ase. Finally, the biochemical energy 3918 is harnessed by a mediator 3920 to produce energy 3922 for a particular process. For example, since proteorhodopsin functions as a light driven proton pump, it generates energy in the form of a proton motive force across the host cell membrane upon illumination. This light-driven proton motive force can be converted to many other forms of energy, one example above being the regeneration of adenosine triphosphate (ATP), via a proton-translocating ATPase. This coupling of the proton motive force generated by proteorhodopsin, for use by proton-translocating ATPases to synthesize ATP, could be accomplished both in living cells, as well as in artificially constructed membrane systems such as liposomes. Proteorhodopsin-based systems can convert light energy to a wide variety of useful mechanical, chemical, and electrical energy forms, for many industrial and technological applications. These

include, but are not limited to, use in targeted drug delivery, uses as primary or secondary energy generators for biocatalytic reactors, fuel cells and nano-machines (including molecular motors), as well as uses in molecular switching or data storage devices.

Applications that can potentially benefit from proteorhodopsin-light driven energy generation are, for instance, bio-electronics applications that are aimed to interface, integrate, or substitute the silicon based microelectronics systems as well as molecular devices. Other applications that can potentially benefit from proteorhodopsin-light driven energy generation are, for instance, in bio-materials, wherein proteorhodopsin is integrated as a bio-material in, for instance, optical films for light mediated computer memory applications, optical information storage and pattern recognition.

Alternatively, proteorhodopsin is useful for a process to enhance yield or increase the potential of recombinant protein production or converting the light induced membrane potential into cellular signals, including modulation of gene expression. The biochemical energy derived from functional proteorhodopsin exposed to light could be harnessed to support a variety of cellular processes. For instance, the energy derived from light-mediated proton pumping could be used to enhance the production of secondary metabolites, or recombinant proteins in host cells, such as *E. coli*. Often, production of specific compounds in the biotechnology industry is limited, since their optimal expression or production occurs in the late stationary phase of growth, when energy reserves of the host cells are low. Retinal-bound proteorhodopsin expressed in such cells would provide an ample source of biochemical energy, by simple illumination. Proteorhodopsin-mediated light driven proton production could enhance any variety of biosynthetic or physiological processes which require energy.

The biochemical energy derived from proteorhodopsin light driven proton pumping could also be converted to other generally useful energy forms, for example electricity. Microbial fuel cells currently use carbon-based compounds, such as glucose, as the primary energy source. Via specific mediators of reduction potential (e.g. electrons), these microbial fuel cells convert cellular biochemical energy to electrical potential. Unlike carbon-based microbial fuel cells, proteorhodopsin uses light as the energy source, that can then be converted into a chemiosmotic potential, and finally into cellular biochemical energy by membrane-bound proton ATP-ases. Therefore, the use of proteorhodopsin could be employed to derive energy from light as the primary or supplementary energy source, that could then be converted into electrical potential (analogous microbial fuel cells that derive their energy from glucose).

In addition to energy generation in vivo in living cells, membranes containing proteorhodopsin could be used to enhance or enable other specific processes in vitro. Polymers produced from proteorhodopsin-containing membranes may have specific properties that could be used similarly to those containing bacteriorhodopsin. One example includes the use of these light sensitive molecules for optical computing applications.

As shown in FIG. 39, the kinetics of proteorhodopsin as it is utilized in 3900 is influenced by various factors such as the type of light source 3914 and the manipulation of light source 3914 in terms of frequency and/or wavelength at which the light 3916 is delivered. Light source 3914 could be any type of light source that delivers light energy 3916 that would be absorbed by light absorbing pigment 3918.

For example, the light source 3914 could be tuned to optimally excite rhodopsin variances with an absorbance maximum of 490 nm or alternatively those rhodopsins with an absorbance maximum of 520 nm. Manipulation of the light source 3914 or the light 3916 being emitted by the light source 3914, for example, involves changing the frequency of fast-light pulses or the delivery of light 3916 as individual pulses, a train of pulses, or a continuous source of light. Manipulation also involves changing the wavelength of the delivery of light 3916 at different wavelengths. In addition, as is clear for one skilled in the art, changing the frequency and/or amount of retinal that will bind within integrated cell membrane protein 3908 also varies the function of proteorhodopsin. Finally, as was mentioned in the previous section, genetic variants of proteorhodopsin result in variants of the proteorhodopsin proteins that changes the kinetics of 3600 due to a difference in absorption of light at different wavelengths. The functional expression of such variation in these proteorhodopsin proteins adds another source of variation to the kinetics of proteorhodopsin as it is utilized in 3900.

As shown in FIG. 39, the light-driven energy generator includes a host 3902. In the present invention, as a preferred embodiment, host 3902 is a cell membrane preparation of *E. coli*. However, the present invention is not limited to the use of *E. coli* and, alternatively, other bacteria or eukaryotes could be used to provide host 3902 as an intact cell (in vivo) and/or as a cell membrane preparation (in vitro). For example, but not limited to, bacteria and yeast with developed genetic systems such as *Bacillus* spp. Species, *Saccharomyces* spp., *Streptomyces* spp. or *Pichia* spp. could be used as host for the expression of proteorhodopsin. In addition, in case a cell membrane preparation (in vitro) is used, host 3902 becomes equivalent to cell membrane 3904.

The light-driven energy generator **3900**, as shown in **FIG. 39**, further includes proteorhodopsin **3906**. Proteorhodopsin is presented in the form of the earlier presented expression vector containing a proteorhodopsin gene or one of its variants. Once proteorhodopsin **3906** has been put into host **3902**, the proteorhodopsin expression vector expresses the proteorhodopsin protein in host **3902**. An integral cell membrane protein **3908** is created in which the proteorhodopsin protein inserts into and folds properly within the cell membrane **3904**. This is accomplished in the *E. coli* host by virtue of the native signal sequence found in the 5' end of the proteorhodopsin gene. It could also be accomplished by replacement of native sequence with another host-specific signal sequence in non-*E. coli* host systems.

As shown in **FIG. 39**, once retinal **3910** is added to cell membrane **3904**, retinal **3910** binds within integrated cell membrane protein **3908** and forms a light absorbing pigment **3912**. The particular example of **FIG. 40** shows an integrated proteorhodopsin protein **3908** bound to retinal **3910** in *E. coli*. Chemical derivatives of retinal (as discussed above) could also be used as a substitute chromophore to generate functional proteorhodopsin. For the particular example of **FIG. 40**, the proteorhodopsin protein was cloned with its native signal sequence and included an addition of the V5 epitope, and a polyhistidine tail in the C-terminus. The proteorhodopsin protein was expressed in host **3902**, i.e. *E. coli* outer-membrane protease-deficient strain UT5600, and induced with 0.2 % arabinose for 3 hours. Cell membranes **3904** were prepared and resuspended in 50 mM Tris-Cl (pH 8.0) and 5 mM MgCl₂. **FIG. 40** shows a proteorhodopsin-expressing *E. coli* cell suspension. After 3 hours of induction in the presence of 10 µM all-trans retinal, cells expressing the protein acquire a reddish pigmentation as indicated by **4010** and the + (plus) symbol. **FIG. 40** also shows that a cell suspension using the same PCR primers

(Sequence ID No:2 and 3) but now in opposite orientation as a negative control, did not acquire a reddish pigmentation as indicated by 4020 and the – (minus) symbol.

FIG. 41 shows an exemplary embodiment of the absorption spectra of light absorbing pigment 3912 upon illumination with light source 3914 as is shown in **FIG 39**. As mentioned above, the light absorbing pigment is a retinal-reconstituted proteorhodopsin in *E. coli*. **FIG. 41** shows absorption spectra of light absorbing pigment 3912 as well as a negative control. After retinal 3910 addition to integrated proteorhodopsin protein 3908, light absorbing pigment 3912 was made. The retinal 3910 addition was done at selected time points, i.e. 10, 20, 30 and 40 min, and shows a progression from low to high absorption values indicated by respectively 4110, 4120, 4130 and 4140 upon illumination with light source 3914. **FIG. 41** also shows the absorption spectra of retinal 3910 addition at these similar time points but now to a negative control of retinal 3910 containing a proteorhodopsin 3906 that was created using the same PCR primers in opposite orientation. 4150, 4160, 4170 and 4180 indicate the four absorption spectra for the negative control. An absorption peak at 520 nm was observed after 10 minutes (4110) of incubation as illustrated in **FIG. 41**. On further addition of retinal, the peak at 520 nm increased, and had a ~100 nm half bandwidth. The 520 nm absorption peak was generated only in membranes containing proteorhodopsin 3906, and only in the presence of retinal 3910. The red shifted λ_{max} of retinal ($\lambda_{\text{max}} = 370$ nm in the free state) is indicative of a protonated Schiff base covalent linkage of retinal to proteorhodopsin.

FIG. 42 shows an exemplary embodiment of the light mediated proton pump of the light-driven energy generator 3900 indicating the conversion of light energy 3916 as shown in **FIG. 39**. The proton pump action is illustrated by measuring pH changes in

the medium surrounding the host 3902, which in this particular example involves a cell suspension of *E. coli*, illuminated by light source 3914. The beginning and cessation of illumination (with yellow light >485 nm delivered by 3916) is indicated 4110 ("ON") and 4120 ("OFF") respectively. The cells were suspended in 10 mM NaCl, 10 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ and 100 μM CaCl_2 . Net outward transport of protons was observed solely in proteorhodopsin-containing *E. coli* cells, and only in the presence of retinal 3910 and light 3916 and is indicated by 4210 in FIG. 42. Light-induced acidification of the medium was completely abolished by the presence of 10 μM of the protonophore CCCP.

FIG. 43 is an exemplary embodiment showing that illumination by light source 3914 generates an electrical potential at the membrane 3904 in proteorhodopsin-containing right-side-out membrane vesicles, in the presence of retinal 3910, reaching -90 mV after 2 minutes from light 3916 onset. Transport of $[^3\text{H}]\text{TPP}^+$ in *E. coli* right-side-out vesicles containing expressed proteorhodopsin, reconstituted with (4310 and 4320) or without (4330 and 4340) 10 μM retinal 3910 in the presence of light (4310 and 4330) delivered by the light source 3914 or in the dark (4320 and 4340). FIG. 43 shows that proteorhodopsin, in its form of 3912 as a light absorbing pigment, pumps protons from the inside to the outside of cell membrane in a physiologically relevant range. The ability of proteorhodopsin to generate a physiologically significant membrane potential, even when heterologously expressed in nonnative membranes, is consistent with the proton pumping function for proteorhodopsin in the native gamma proteobacteria from which it is derived.

FIG. 44 is an exemplary embodiment showing that proteorhodopsin can have a fast photocycle and can therefore be characterized as a fast and therefore efficient

transporter of protons. For the particular example of FIG. 44, light absorbing pigment 3912 is induced by laser pulses delivered by light source 3914. Laser pulse-induced absorption changes are shown by 3912 in host 3902, which in this case are suspensions of *E. coli* membranes containing proteorhodopsin. A 532-nm pulse (6 ns duration, 40 mJ) was delivered at time 0 and absorption changes were monitored at various wavelengths in the visible range in a lab-constructed pulse photolysis system. 64 transients were collected for each wavelength. 4410 indicates transients at 3 wavelengths exhibiting maximal amplitudes. 4420 indicates absorption difference spectra calculated from amplitudes at 0.5 ms (indicated by 4430) and between 0.5 ms and 5.0 ms (indicated by 4440). In 4410, transient depletion occurred near the absorption maximum of pigment 3912 (500-nm trace indicated by 4450), and transient absorption increase was detected at 400 nm (indicated by 4460) and 590 nm (indicated by 4470), indicating a functional photocyclic reaction pathway. In 4420, the absorption difference spectrum shows that within 0.5 ms an intermediate with maximal absorption near 400 nm is produced (indicated by 4430), typical of unprotonated Schiff base forms (M intermediates) of retinylidene pigments. The 5-ms minus 0.5-ms difference spectrum 4440 shows that following M decay an intermediate species red-shifted from the unphotolyzed 520-nm state appears. The decay of proteorhodopsin final intermediate is the rate limiting step in the photocycle and is fit well by a single exponential process of 15 ms, with an upward baseline shift of 13% of the initial amplitude.

As mentioned above, a proteorhodopsin gene or protein variant can be selected to determine an absorption spectra of the light absorbing pigment to change the kinetics of the light energy generator 3900, for instance to meet a design/functional criteria of an application wherein proteorhodopsin is utilized. FIG. 45 shows an exemplary

embodiment of different absorption spectra of retinal-reconstituted proteorhodopsins in *E. coli* as a function of wavelength 4510. As shown in FIG. 45, the absorbance 4520 is different and depends on the clone from which the proteorhodopsin was amplified. In this particular example, 5 μ m all-*trans* retinal was added to the membranes suspensions in a 100 mM phosphate buffer, with a pH 7.0, and absorption spectra were recorded. The four spectra 4530, 4540, 4550, and 4560 are respectively for the proteorhodopsin genes retrieved from clones HOT75m4, PALE6, HOT0m1, and BAC31A8 at 1 hour after retinal addition. The proteorhodopsin gene retrieved from clone HOT75m4 4530 and PALE6 4540 produced a blue (490 nm) absorption maximum. The proteorhodopsin gene retrieved from clone HOT0m1 4550 and BAC31A8 4560 produced a green (527 nm) absorption maximum. In general, a range of wavelengths could be obtained that is not limited to the range shown in the example of FIG. 45.

It will be clear to one skilled in the art that the above embodiment may be altered in many ways without departing from the scope of the invention, such as for instance by mutagenesis to change the genetic sequence of proteorhodopsin and thereby changing the kinetics of the proteorhodopsin protein once it is expressed. Accordingly, the following claims and their legal equivalents should determine the scope of the invention.

DEPOSITS

Depository address: 10801 University Boulevard, Manassas, VA 20110, USA.

The *Escherichia coli* containing cloned DNA BAC 31A8 having assigned ATCC number PTA-3083, the *Escherichia coli* containing cloned DNA BAC 40E8 having assigned ATCC number PTA-3082, the *Escherichia coli* containing cloned DNA BAC 41B4 having assigned ATCC number PTA-3080, and the *Escherichia coli* containing cloned DNA BAC 64A5 having assigned ATCC number PTA-3081, all having been deposited on February 21, 2001 with the ATCC Patent Depository.

The *Escherichia coli* containing a plasmid PAL E6 having assigned ATCC number PTA-3250, the *Escherichia coli* containing a plasmid HOT 0m1 having assigned ATCC number PTA-3251, the *Escherichia coli* containing a plasmid HOT 75m4 having assigned ATCC number PTA-3252, and the *Escherichia coli* containing cloned DNA BAC64A5 having assigned ATCC number PTA 3082, all having been deposited on March 30, 2001 with the ATCC Patent Depository.



LIGHT-DRIVEN ENERGY GENERATION USING PROTEORHODOPSIN

LIST OF SEQUENCES THAT ARE LISTED IN THE INCORPORATED SEQUENCE LISTING

- Sequence ID No:1** bacterial artificial chromosome (BAC) clone 31A8 (EBAC31A8).
- Sequence ID No:2** nucleotide sequence of proteorhodopsin-specific polymerase chain reaction (PCR) primer 1.
- Sequence ID No:3** nucleotide sequence of proteorhodopsin-specific polymerase chain reaction (PCR) primer 2.
- Sequence ID No:4** nucleotide sequence of the proteorhodopsin gene amplified from clone EBAC31A8 (Sequence ID No. 1) using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:5** deduced amino acid sequences of the proteorhodopsin gene amplified from clone EBAC31A8 (Sequence ID NO:4).
- Sequence ID No:6** native proteorhodopsin nucleotide sequence from clone EBAC31A8 (Sequence ID No:1).

- Sequence ID No:7** deduced amino acid sequences of the native proteorhodopsin nucleotide sequence from clone EBAC31A8 (Sequence ID No:6).
- Sequence ID No:8** nucleotide sequence of the proteorhodopsin gene amplified from clone EBAC40E8 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:9** deduced amino acid sequences of the proteorhodopsin gene amplified from clone EBAC40E8 (Sequence ID NO:8).
- Sequence ID No:10** nucleotide sequence of the proteorhodopsin gene amplified from clone EBAC41B4 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:11** deduced amino acid sequences of the proteorhodopsin gene amplified from clone EBAC41B4 (Sequence ID NO:10).
- Sequence ID No:12** nucleotide sequence of the proteorhodopsin gene amplified from clone EBAC64A5 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:13** deduced amino acid sequences of the proteorhodopsin gene amplified from clone EBAC64A5 (Sequence ID NO:12).

- Sequence ID No:14** nucleotide sequence of the proteorhodopsin gene amplified from clone HOT0m1 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:15** deduced amino acid sequences of the proteorhodopsin gene amplified from clone HOT0m1 (Sequence ID NO:14).
- Sequence ID No:16** nucleotide sequence of the proteorhodopsin gene amplified from clone HOT75m1 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:17** deduced amino acid sequences of the proteorhodopsin gene amplified from clone HOT75m1 (Sequence ID NO:16).
- Sequence ID No:18** nucleotide sequence of the proteorhodopsin gene amplified from clone HOT75m3 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:19** deduced amino acid sequences of the proteorhodopsin gene amplified from clone HOT75m3 (Sequence ID NO:18).
- Sequence ID No:20** nucleotide sequence of the proteorhodopsin gene amplified from clone HOT75m4 using PCR primers according to Sequence ID No:2 and Sequence No:3.

- Sequence ID No:21** deduced amino acid sequences of the proteorhodopsin gene amplified from clone HOT75m4 (Sequence ID NO:20).
- Sequence ID No:22** nucleotide sequence of the proteorhodopsin gene amplified from clone HOT75m8 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:23** deduced amino acid sequences of the proteorhodopsin gene amplified from clone HOT75m8 (Sequence ID NO:22).
- Sequence ID No:24** nucleotide sequence of the proteorhodopsin gene amplified from clone MB0m1 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:25** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB0m1 (Sequence ID NO:24).
- Sequence ID No:26** nucleotide sequence of the proteorhodopsin gene amplified from clone MB0m2 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:27** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB0m2 (Sequence ID NO:26).

- Sequence ID No:28** nucleotide sequence of the proteorhodopsin gene amplified from clone MB20m2 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:29** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB20m2 (Sequence ID NO:28).
- Sequence ID No:30** nucleotide sequence of the proteorhodopsin gene amplified from clone MB20m5 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:31** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB20m5 (Sequence ID NO:30).
- Sequence ID No:32** nucleotide sequence of the proteorhodopsin gene amplified from clone MB20m12 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:33** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB20m12 (Sequence ID NO:32).
- Sequence ID No:34** nucleotide sequence of the proteorhodopsin gene amplified from clone MB40m1 using PCR primers according to Sequence ID No:2 and Sequence No:3.

- Sequence ID No:35** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB40m1 (Sequence ID NO:34).
- Sequence ID No:36** nucleotide sequence of the proteorhodopsin gene amplified from clone MB40m5 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:37** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB40m5 (Sequence ID NO:36).
- Sequence ID No:38** nucleotide sequence of the proteorhodopsin gene amplified from clone MB40m12 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:39** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB40m12 (Sequence ID NO:38).
- Sequence ID No:40** nucleotide sequence of the proteorhodopsin gene amplified from clone MB100m5 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:41** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB100m5 (Sequence ID NO:40).

- Sequence ID No:42** nucleotide sequence of the proteorhodopsin gene amplified from clone MB100m7 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:43** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB100m7 (Sequence ID NO:42).
- Sequence ID No:44** nucleotide sequence of the proteorhodopsin gene amplified from clone MB100m9 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:45** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB100m9 (Sequence ID NO:44).
- Sequence ID No:46** nucleotide sequence of the proteorhodopsin gene amplified from clone MB100m10 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:47** deduced amino acid sequences of the proteorhodopsin gene amplified from clone MB100m10 (Sequence ID NO:46).
- Sequence ID No:48** nucleotide sequence of the proteorhodopsin gene amplified from clone PALB1 using PCR primers according to Sequence ID No:2 and Sequence No:3.

- Sequence ID No:49** deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALB1 (Sequence ID NO:48).
- Sequence ID No:50** nucleotide sequence of the proteorhodopsin gene amplified from clone PALB2 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:51** deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALB2 (Sequence ID NO:50).
- Sequence ID No:52** nucleotide sequence of the proteorhodopsin gene amplified from clone PALB5 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:53** deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALB5 (Sequence ID NO:52).
- Sequence ID No:54** nucleotide sequence of the proteorhodopsin gene amplified from clone PALB7 using PCR primers according to Sequence ID No:2 and Sequence No:3.
- Sequence ID No:55** deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALB7 (Sequence ID NO:54).

Sequence ID No:56 nucleotide sequence of the proteorhodopsin gene amplified from clone PALB6 using PCR primers according to Sequence ID No:2 and Sequence No:3.

Sequence ID No:57 deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALB6 (Sequence ID NO:56).

Sequence ID No:58 nucleotide sequence of the proteorhodopsin gene amplified from clone PALB8 using PCR primers according to Sequence ID No:2 and Sequence No:3.

Sequence ID No:59 deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALB8 (Sequence ID NO:58).

Sequence ID No:60 nucleotide sequence of the proteorhodopsin gene amplified from clone PALE1 using PCR primers according to Sequence ID No:2 and Sequence No:3.

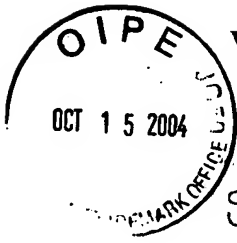
Sequence ID No:61 deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALE1 (Sequence ID NO:60).

Sequence ID No:62 nucleotide sequence of the proteorhodopsin gene amplified from clone PALE6 using PCR primers according to Sequence ID No:2 and Sequence No:3.

Sequence ID No:63 deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALE6 (Sequence ID NO:62).

Sequence ID No:64 nucleotide sequence of the proteorhodopsin gene amplified from clone PALE7 using PCR primers according to Sequence ID No:2 and Sequence No:3.

Sequence ID No:65 deduced amino acid sequences of the proteorhodopsin gene amplified from clone PALE7 (Sequence ID NO:64).



ttgttatatc agtaatggct attgctccaa taacttaata ctaatatata attagtttat 60
gaataaaattt tatatatattg ggttatgtt ttttacta aatgcatttt cttgctcaga 120
tcttctagat acagacatga gagttcttga ttccgctgag tcaagaaaacc tttgcgagtt 180
tgaaggaaaa gctttactag ttgtgaatgt tgcaagtaga tgtggttaca cttatcaata 240
tgctggcctt caaaagttat atgaaaagtta taaagatgaa gatttttctag taattgggat 300
cccatctaga gatttttcttc aagaatactc tgatgaaagc gatgttcgag aattttgttc 360
tacagaatac ggtgttgaat ttcctatgtt ctcaactgct aaagtcaaag gaaaaaaagc 420
acaccattt tataaaaaac ttattgcaga atcaggtttt actccctcat ggaactttaa 480
taaatactta atctcaaaag agggcaaggt tgtatccaca tatggatcaa aggtaaagcc 540
tgattcaaaa gagcttatat cagctataga aggcttgctg taaaattatt acttagaaac 600
taatacagtt ttaggcttgt ttgctgcaaa tattccatta tctacaactc caggaatat 660
attaatcaaa gcttccattt cagtgggggt tgaaatatcc atattagaga tatctaaaat 720
gtgattacct tggctctgtta taaatccagt tctatatgtg ggtattccac cgatcgagat 780
tatttttctt gcaacaaggc tcctacttcc aggtatcacc tctataggca gtggaaaaagc 840

tcccaaaaga ttaaccatct ttgactgac aactatacat ataaactcgt tagaggcaga 900
agcaactatc ttttctctag tatgtgcgc accaccacct ttaataagac aattttcagg 960
agacacctca tctgcaccat ctatgtaata agctatatca actacatcat taaggctaaa 1020
gacctctatc ccattttcat ttaataattt tgatgaagca tctgaactag aacagctcc 1080
agcaaatctg tgcctatgct ccttagttc ttctataaaa aaattaactg ttgagccggt 1140
tccaatacct aaaatcatct caggatgaag attatttttg atatatctta tagcttggtt 1200
agcaacattt atctttgagc cactcataga gttataatac aagaaaaatat aggtagttaa 1260
ttattttgag actaaaaaatt aaaaaaacag gttcttttaa gaattcccag aagtacctaa 1320
agcttattcc taatgcatct gtttatgacg tagcattaaa atcacctata acatttgctc 1380
taaataattc ttcaaagctg gggaaataaag ttttcctaaa aagagaggat ctgcaacctc 1440
tattttcttt taaaaacaga ggagcgtata acaagattgt aaatttatcc gatgccgaaa 1500
agaagagggg ggttattgct gcatacagcag gaaatcatgc tcaaggggta gccagtgc 1560
gtaagaaaatt aaaaattaat tgcttgatag ttatgccaat acaactcca gaaataaaaa 1620
taaaagatgt aaaagattt ggagccaaaa tactccaaca tggggacaac gtagatgcag 1680

cattaaaaga ggcactgttt attgcaaaga aaaaaaatt gtcttttggtt catccttttg 1740
acgaccctct aacaattgct ggccaaggga ctataggaca agaaatttctt gaagataaaa 1800
ataattttga tgttgtcttt gtccgggtgg gaggaggagg tattctagct ggtgtatctg 1860
cctggatagc acagaataat aagaaaataa aaattgttgg tgttgagggtt gaggattccg 1920
cttgtcttgc tgaggccgta aaagctaata aaagagtttat tttaaaaagaa gtgggcctct 1980
ttgtctgatgg ggtggcagta tcaaggggttg gaaaaaataa ttttgatgtt attaaagagt 2040
gcgtagatga agtcattaca gttagcgttg atgaggctctg caccgctgta aaagatatct 2100
ttgaagatac aagggttcta tcagaaacctg ctggggcatt agcacttgca gggttaaaag 2160
cctacgcaag gaaagttaaa aataaaaaac ttattgctat aagttctggc gctaattgtaa 2220
atttcccaag acttaatttt attgttgagc gatcagagat tggtgaaaaat agagaaaaaa 2280
tattaagtat caaaatccca gagataacctg gaagttttct taagctttca aggatgtttg 2340
gcagctctca agttacagag tttaaactaca ggaaatctag cttaagcgat gcatatgttt 2400
tagttggtgt tagaactaaa actgaaaaat catttgaaat cttaaagtcc aaattaaaaa 2460
aagcaggctt cacctttagc gactttactc gaaatgaaat atccaatgat catctgaggc 2520

atatggttgg tggcagaaat agtgactcag gctctcataa caatgaaaga atatttaggg 2580
gagagtttcc tgagaagccg ggcgcgctgt taaattttct agagaaattt ggaaataaat 2640
ggmatatttc cttatttcac tacaggaacc taggttcagc ttttgaaag atattaattg 2700
gcatcgagag taaggataaa gacaagctaa taaatcattt aaataagtca ggnactattt 2760
ttacagaaga aacctctaac aaggcataca aagattttttt aaatgaaag gttaatactt 2820
taatctaaat ttaattgaaa aaagctcatc gctagggtttt tcccacggct ctttgaacaa 2880
ctcggattga gatctatcat cctcctcgtc gtaaattctc ccacctttag aatagaccaa 2940
aaatagatat gacaaaggag cgagctcata tttatatcta atttggaacg aagctacgcc 3000
agtattaaat tcattaacta tattatttcc tttataaagg tatccattag catctgaaaa 3060
aatactaata ggatttttttg ctttcaaagc aacaaaattga ctcttaagtc tgatctcatg 3120
tttattattt ttaaaccaat ttagatcaaa agataaggta tcttggtcttg aatcatatga 3180
ggcaagatta ttattatcct gccatatcag ccatttcattt tccttttctta ttctatatg 3240
cgcattgatt cttaagttat catttggaag tattgaaacct gctatcttgt aaaattttct 3300
accataccca ttcgaatccc accgattatc tttctctccc ttaagaagc taactctcca 3360

gtcatatgtc cagaatgaat agttcttttg ctcaaagtct gctgtaatac ctattcttct 3420
ttttgatttg ataaaggggt aggcttcatt ttttcttgat atagttgtat ttttcccaga 3480
agatctaaag ttaaaatcta attgaaattt agagttgtcc ttaaaactaa aagaattttt 3540
ttgatcgaat cctattggat ttgaattacc agctgtgtca gcatcataat ttagatcaat 3600
tccataatct atttgtttta atatcgagct attatcaaat tcattttattt ttcgattttcc 3660
tccaatacca gcatgaatcc agtctcttct ttgcagataa ccaaagtcac ttaactcgaa 3720
gtcgtcttca aaataaagaa ggcttccact tatgttagat agtttatttg gaagatatgt 3780
aaactgagtc ctatacccaa gccattttt accatctttt tctgaggcta acagatctga 3840
atatgtaatt aatttttttg aacgaatatt gatgtaatca ataacattga ctggtgatga 3900
ttcgcctgtc atttcattct caacattcgt taccatgaag ccaagcgttt tatttccaag 3960
ctttgttcga gatcgaagag cataatagtc tcttccaact gaaaaggctt catcagcctc 4020
acttgctaca aatactccaa attcattatt attacttttt tgagtaagtc ttaatgcaa 4080
atcaatatca gaatagtttt ttttagctgc ctgcgaacct tcttcattac tttcttctga 4140
gcaattatag ctgggggcag ctccaatcct ccgcgtattt ataaccgagt atctatcata 4200

attactaata tcaaatagtg attggttttc attgaaaaat gctctttttt ctgagtaaaa 4260
agtttcttga gcagaaaaagt taataaccac atcatcactc tcagcttgtc cgaaatcttg 4320
attaatagct aaattttattt gacgaccttt tccagtgcta taaaagattt cagccccaat 4380
atctgaccct tcttggttttg taactgaatt tttattttgaa gatataatg gaaaaaaagt 4440
aagctttgat tttgtatagt tttgtatttc taagctctct aactcttgaa agtagtcatt 4500
tctactagcc attgttcogg cactgctaac ccatgactca ttttcggcac tataacgtaa 4560
tgccggtgtaa ttaatttttc ttatacacc atcaggctgt ttcatatacg ttacatccca 4620
aggaataaaa aactcagaga cccaataccc atcaaaatttt tgtgtttttg caatccaatc 4680
tccatcccg tctgtcttaa agtctcctgc ttgcgttttt atggcatcga aaagcgagtt 4740
cccaagattt atagcaagaa tgaaagcttt gttaccatca ccatcaaagt ctatatatt 4800
agagttttta tcgctaagtg aattttatttg atctctaagc gtcttcctgg agaacaataga 4860
atcattactt tgaaaattct taaatccaac atatatacca tccttatttg agaaaatata 4920
agccggttga agaagtcat ttttcttaag agtaaaagga gatgtttcat aaaaatctgt 4980
aatttcaaat gcattattcc actcaggctc atcaagagag ccatcaataa caattgagtt 5040

tgaccaaatc agcacagagg taagtaaaag tgataatgag gctaaagtta ttttcataga 5100
tagatttaat tcagagtatt ttaatctatg aagagtagga aatcatccga tcattctcag 5160
aaaaaacata atcatattta aggtcatggt tttctccaaa atgcataatta caaatctggt 5220
attgataaca cagtccaacc aataaaggtc ttgatacctc ctcgtttata gagccaataa 5280
ttctatcgaa atatccagag ccatagccaa gcctgtatcc atttaaatca actcctgtca 5340
taggaataaa cattaaatca atttcattta tgttgacata atcttcactt ttaacctctt 5400
tgatcccaaa ttgatttata aagaagttag gctgctcatc cagcaaatca aaagccatca 5460
tttcgtcatc aattactttt ggaatataaa tattcttttt aagttagta aaagcttgaa 5520
ttaacaaatt tgtattgact tcatttcgaa aaggaaaata taaggcaatg ttttgcattt 5580
catgagtatt gattttttct aatacatttt cttaggattaa aaaacttata ttgctcttag 5640
ataaatcaga aatagactgc ccctgctcaa agagtgattt tcgtatttta tttttcacca 5700
taaatgtgac cgaggctaac aaaccaaata accgccatgt gctataacctg aaccgatggg 5760
tcagggtgtg aatttcataa caaatcaggc ttccctgtaa aggtactgag caacaatgct 5820
agaggaatat cacctattta attgtatcgg ttcaaattta ttaacacatt agcgtatgaa 5880

ccagaatgca gttaattata aaatatatag gtattaagta aaagttaatt tttagagagc 5940
agactctatt ttttgtatta gcttttcgat atctttatta tttaaagcag agtcagtatt 6000
ctcgggtggc ttggataaaa gcttactagc catagttaga ccagctatca ccaaggcatt 6060
attcttatca ctgatccat caagctcatc atttaataat ttgagcctc taatgagttg 6120
gtctttctct tctggcggac aggctaactc cagatctcgg ccaaaaattc ttaaagatag 6180
cgtttccatt tctgacatta ttttttgagt ttagttattt ctttatttag tatgagctta 6240
tcttttttcc agtctctttc cctttgcaca taagaatcta ttattccttt ttgctcatca 6300
aatcttttta tcaaaagatc taccttatcc tctagttcaa ataatgatga ttttttttct 6360
tctgacattg tataagtta attgtactag aatgaattgg aaagtttctt taataaaaga 6420
gtaaaataag gcatggaaaa aataattttt aaaaacagaa gagattcact tataaaacac 6480
ctacctaaaga attcagcctt aattgtacct ggtgcagatt tgcaatatag aaatgctgac 6540
tcactttata atttaagaca ggaaagtagc ttctactatc tgtctggcctt ttgtgagccc 6600
tcttctctaa tggtttttagt taataatgga aaaagcattg attcaataat ttttgttcct 6660
gaaaaagata aacttaaaga aatctgggat ggttatcgag ctggccctga gggcgcaata 6720

aatgattttc tttttgatca agctttttgaa aataataaat cagatgcttt aatgcctgaa 6780
atccttcaag ggctagaaaa agtttttttat tcaataggga agaaaaatgg ctttgatcag 6840
aaagtaattg actggacatg cgcagcaaat tctaaagata ggcacagcaa atcaattgat 6900
attattgatg gctcttcgat ggtaggaaat ttaaggctta tcaaagataa gcatgaaatt 6960
gatatatga agagagcttg tgaaatttca gctgaatcat atattgaggt catgaaatct 7020
ataaagcctg gggacaatga gcaggaaata gaggcgctat ttttatatga attcgccaaa 7080
agggaggaa ggtttccagc ttatacacct atagttgctg gaggtgaggg tgcttggtga 7140
ttgcattata ttgaaaaatga taaagagtta gcttcatcag atttaatttt ggtagacgca 7200
ggatgtgaat acaaaatgta tgcattctgat atcacagaa cttcccagt aagtggaaaa 7260
tttccagatg aacagctaca aattataat atgtccaca agccaatct tgctgcaatc 7320
gatgctgtaa aaactggtaa tagcataatg gagcccaaa tggtttcaga aaaagtaatt 7380
actgaaggtc ttgtagagtt gggatttcta tctggcgatg ttaatcagct tcataaaaat 7440
ggtgcattca aggactttta tatgcataag gtgggacatt ggcttgact tgatgttcat 7500
gatgttggtg actacatgga gggagatgag tttatgaagt ttaagccagg gatgataacc 7560

acaatcgagc caggcatcta tatcagtagc gcaatggatg tagatgacaa atggaagagg 7620
atcggcataa gaatagagga cgacatcctt gtaacagatt caggcaatat taatctaaca 7680
gagaagggtgc catctaatac tcaagaaata gaatcattga tggcttagac tatggagggtt 7740
ccaattgtta tttctggcgg agggataata ggtaattaca tttctcttag gcttgaaaaa 7800
aataatatca aaaccgttat tgtcgaaaaa gctagtagtt tcaaagccct agataagggt 7860
ataagaacag tcactctcaa tgagcattct atgcaaatgc taaaaaatat tggattttgc 7920
ccatcaattg ctcaaatcaa cagcatcgac gtattagatg gtgagggtac aggcaaaatt 7980
caatttctag caaaggacgt aggcagcgaa aacctttcat atgtaacctt tttcaatgaa 8040
ttacaaaaac taatttctga tccatgtaaa gaaagaacct tatttaataa tgagattgat 8100
tcagttcaga atcttaatac agaattctgat ccagagatca tgcttaaaga tggcatgacc 8160
ataaaaaacga atctaattgc tggatgtgat ggaagaaatt caaatattgc aaaaattgct 8220
tcacttacaa gcagcttcga tgactactta caaacagctt taacttttgt cgttgatatt 8280
gataatgatt cacatggcaa agctcaccaa gtttttctg aaaaagggaat atttgcactt 8340
atgccactcc cagaaggcaa ggtgagatg aataaatgca cagtggtttg gtcaataaaa 8400

aatcaagttt tgggagatga gcctgtatct gagtttgtaa aaaataacat ttcttttttt 8460
gaatcaaagc ttaatgttag tctcagggtt aagtcagaaa ttttaagttt taaattatcg 8520
aaccatcatt ttgaaaacta tattagcggg cctattgttc ttcttgggtga tgctgctcac 8580
tcaattcacc ccttagcagg tcaagggtatt aatctaggat ttgcagatgc agatactttt 8640
tgtgaagagg taattagttc ttataaaaaa gggattgcct ttaatgagaa atcagtttta 8700
aaaagatatg agattagaag aaaaagtatg aacttttttaa tgttgaagtc tatggacttt 8760
tttgtggatt tatttggttc agaaaattta tatctaaggc tgatcagaaa tttaggtatt 8820
tcttcgttaa ataaatcaaa gtttgttaaa gcatttcttta taagacatgc ctctggaatg 8880
aataagtttt aaatttgtat taaacttttt gaccttttagc tctaagttct ttàagaactt 8940
cactaatgcc ttttttatca atgattctca taccttttgc agatacttta agatttacga 9000
acctgttctc agattcaacc caaaatttgt gtgtgtgaag attaggaaa aactttcttt 9060
tagtcctatt tttagcgtga gaaacattgt ttctgactg tggatatctta cctgttactt 9120
gacataattt actcattgaa acgcgatttt atagaacact gaggaactta gcaatactat 9180
tgtgaaacaa atttatttat tacggcatgc acaatctgac tgggagagct ctaatcagaa 9240

agattttgat agaccattag caagaaaagg cattgaagaa gcaataaaaa tatcatgtta 9300
ctgcaaatct cattcaatct tagtagataa aatattctgt agcactgcag aaagaactaa 9360
gcagactttt gatatatgca gtgatgggct taattatcca atagctgaag cagtctatac 9420
tgatgagctt tacttttctg gccctggtga gtcgttaag cttatccaaa gtttaagtga 9480
attcatttcc tctgttttaa taataggcca caatccatca atgcaaatgt acatagatgc 9540
tatttcagaa aatcctcata ttacgtattc aacatgggg ctggcagaaa ttctcgttga 9600
aagttcatgg aaagacttat ctttaaaaa atgtaagta aaatctttaa ttcaaccagg 9660
agagctttaa aaattgaaaa acataaaaa taagatcatt aaccactaa tgggatccaa 9720
gataccctta cctcaatatg aaacaaaagg ctggcagga ttggatttaa gggcatgcct 9780
agatagtaat ctcagccttc aagcaggaa atctcagttg atacctattg gttttgcaat 9840
gtacttagaa gatcccggtc ttgcagcaat ggttatacct agatcagggt taggttctaa 9900
gcatggaatc gttcttggta atctggttgg gttgattgat tcagactatc aaggagagct 9960
aatgggttcct gcctggaata gatcagatac agatttttgag attaatcctg gagacaggat 10020
tgcacaaatg attatagttc cagtgattca agcagatttt gaaattgtag acgagttcaa 10080

tgagactcag aggggagaaaa aggggttttgg aagttcaggt ataaattgat aaatttactt 10140
tttcttgcca aatcttttctt caaatcttctg aactcttccg ccagtatcaa taatttttttg 10200
cttaccggta taaaaaggat gagaagcaga ggatatatca agaggtagt atgggtatgt 10260
ttttccatct tcccatctt tcgtttgagt cgtatctaata gttgaacgaa tgagaaaagaa 10320
cttatcagca ctacggtcat ggaataaaac ttcacggtat tcaggatgta tatcttttttt 10380
cataataaaa tcaaaaatttg gatgagaact ataacaaaaa acaacttctt ttcaatcaaa 10440
acatgaaaaat tttttactat gacatagctg tttcattgcc tctaaggcaa tgctttactt 10500
atagtttctga gcttaaaatt acaaaaggaa cccgcgtgtc agttcctttt ggtaaaaagga 10560
agattgtagg agtgggttata aaaaatatc aaaagccaga tttcctaaaa aaagcggag 10620
ctataaaaaa aattatcgct gtacttgatg aatatccttt gttcgacaaag ccaatttttg 10680
attctatat gtggtcttct gattactatc atcatcctat tggtgagggtt ttaataacct 10740
ttatacccac cgaattaaaga aaaattaata ataaaaaaat tgaagcttta agagaatttt 10800
ctgaatatc agtaaatgag gatgataaga aattcgattt aaccaaggat caagaaaaag 10860
cagtcaaggc gcttttctaaa tctaaaggat tttcaccac tttattatat ggagttacag 10920

ggtctggaaa aacagaagtt tacttaagag ttgccgaac ttttattaaa ataataaagt 10980
cagtattagt tttagttcca gaaataaatt taacccccca attgctctcg cgatttgaga 11040
atagatttaa tggtgaaatt ggcatatata attctaagca aacagcagct aagagattaa 11100
agacttggt aaaagctaaa ttgggttcta taaaaataat agtaggaact cgatcttcgg 11160
ctttagtgcc tttagataac attgggtttaa taattatcga tgaagagcat gaccaatcat 11220
ttaggcagtc agaagggttt aaattctctg ctagagactt aagtataaaa agggcacagc 11280
ttgcagatat tccaattatt ttgggatacag caacccttc gctgcaaaact ttaaaaacttg 11340
taaaagaaaa taaatttata agagttgata ttcctaatacg agttgatgga aacaagcctc 11400
ctaaaattaat agccttagat atcaataaca gccctttaat aggcggagtt gctaaaagaga 11460
caattgaagc aatgcaatca accatagaca gaggagaaca ggttctagtt tttatttaata 11520
gacgaggatt cgctccactc tatcaatgta gtagttgtgg gtgggtagca gattgtataat 11580
cttgtgatac aaatttagtc ttccaccagg caagaaatag attaatattgt catagggtgtg 11640
aatctgccta ctctgttaat ttgtcttctc cggcatagcaa gtctaataac ttaataatgt 11700
atggagctgg aacagagaga gtgaagaag ttcttaaaag cagctttgta aagactccaa 11760

taattagagt tgatcatgac tcaacaaaaa aagtgggagc tatggaggct atagttaaaa 11820
aaattcattc ctcagacgca gcaatttttag ttggaactca aatgcttgca aaaggacatg 11880
attttcctaa agtcaccta agcgttat tttaagtctga taatggcctt ataagcccag 11940
aaattaatgc attagagaaa atatctcaat tgcttattca ggtctctgga agagcaggaa 12000
gaaataataa tcttgcaaaa gttattattc aaacaagata tcctgatgat ataaatctta 12060
ataaaaataa gacaggagat tatatgaaat ttgctttctca atgcctaagt accaatgagc 12120
aatgaaactt acctccattt actactttat gtctgcttag gtgctcatca ccaactcaaa 12180
agagtaatgt agatttccta gagaaagctg ttttaatttt atccaatagg actgatatata 12240
atgttattgg tcccttgcct tcattagttt cgaaagtcgaa aggaaaattat aggcaccaag 12300
tctatatcca tgcaccaaag aagacttttt taaataaggt attaaagttt ttgacaacag 12360
agtttgaaaa atggccggaa tctaataagg ttaagtggtc ttctgacatt gatccaatag 12420
acttaagcta aataattaatc ttaattaatt gtctgggta tattgggtta ttgtttagtt 12480
tattctctgt attaatctct tctacagtca ccccaaatct tctgctatt tctgataaga 12540
catccccctt ttgtattttg taagtcacaa agcctggatc aatactcata aaggtatttg 12600

ctttaggttt gtccttaaaa tagttatgta ttcctaggaa aacagatctt gcaatcattc 12660
ttctccctgg cttgcccttt aatctttgtg cgtcttcagg gttgggtata aaccctgact 12720
caaccaatac agaaggtata tcaatagact taagaactct gaaatcagcg tactcaacat 12780
tctttttatg aatttttgtg aaagggctct ttttaagttg atccaatata ttagttccta 12840
aaattttact ttcagaaaatt tttttcttat atatttctgg ataggtttgt cttgcccgcat 12900
cctcatcaaa atcaactggc ttttagatttt ttatatcagc ctgtattctt tcccttttgtt 12960
tttttagataa gttccttagca acagtacttg aagcttcac tgaccatata aaaacagaag 13020
caccctttac ggaagataat ctaaaccat cagcatggat tgaaacaaaa atatctgctc 13080
catatttct agcatcctgg tatctattat ttaaattctaa tgtctcatct ccatctctaa 13140
tcatcacccg tctgtaccca taagtatctc ttaagggttct ttctaactcc ttcgcaataa 13200
gtaacgtcac atctttttct aaaatatatt ttggggccaac cgcaccaggg tatttaccac 13260
cgtgacccgc atctatagca acaacaatat ctctaatact tttattgagg ttttttatttc 13320
ttttcacctt aagctctaata tttatatatt ctgtatttat agtctgagtt gggtttttgcc 13380
aatggactga ttcataataa tcaacaacga ttctggtaag acttccatcc tgagatgctc 13440

ttactttttt gatggatag ttgtatggaa catttatctc ggtcggaagg ctcgattgat 13500
taatttccat gacaattcta gagggatttt taaagaataa gaccttacca atgaaaacttt 13560
atctagacta aagttaatac taatttcatt gttaccocatg tcttgtatatt catcgaagaa 13620
aacatcattt ccgctgataa aaaagatag aaaaccaagg attgccaatc tattcataag 13680
tttttaacca ttctttaaag ccattgtcac cagaaaattaa agaaaacctct ctcccttcag 13740
gaagatggct gaaaattatt tttagatcaa aacttcttct atgttgaagc ctttcaggcc 13800
actcaattag cactacttct ttagagtta ttttctact tagatcaaat atatcaatat 13860
cttctgcttc gtaggttcta taaagatcaa tatgcaaaaa catataaata ttaaaatcat 13920
attcctcgca gagagtgtaa gttaggacttt ttacaagatc ctcccatcca caatttttta 13980
taatagatct agatataaag gttttgcctg ctctaagtc tccctcaaga tgtattttcaa 14040
tttcttgcga agaagattta agtatctcca tagctatttt tgaccctagt tggttttag 14100
cctcgtcatt tataagagta agtttttttca tcgatttatt aattgtctta ttataggat 14160
taaactagag gcattttaagc ctatttcacc aatatctacc taaaatttta aacctgcctc 14220
cgaatgaaca gcaacagcaa taatgctcgc atttcttata tcaagacctt gggcaagaag 14280

agcagtcaaa acaccagcca gaacgtcacc agttcctcct gaagcaagtt caggtccgcc 14340
gcacgcgat ataaaagact tgttatttgt tgagtcaaag accaccgttc ccattccttt 14400
taatatgaca atagaagccc cataagtgtc tgcaatttgc ttggcagcag aaatccctatc 14460
tctttgaact tcttcaatag atatgttttag taatatagcg gcctctcctg gatggggcgt 14520
catcaagatt gttttattgg atttttttgat taaagatttt gatgatgcca cgatgtgttaa 14580
agcacctgca tctaataataa ttttttgaatt atttgcagac tttaatatatt tacccaatat 14640
agttttttgca aaagcagtat tagcaattcc tgggtccatac aataataacg tatgatcttt 14700
aattttttaa tctatatcga ccccttcaac catatcaaca ccaattgcca ttacctcagg 14760
gtttcttaat aaagatgggc ttacattaga cgtgtctgtg acgagagtta ctagtccaga 14820
gccacaaaat aaacttgcct cagatgcaag tattccggcg ccacccatgc ctgggggagcc 14880
agcgcatatc aaaacttttc caaaattacc tttatggcta tcctttgtctc tattggggcag 14940
aagatttttt aattcttgaa aggtaaaaact ttgcaatata gacatgaaat tacttaagct 15000
aaattatgtt gaatgctaaa aattcattct tcagcgaaat aaccaatcct atcgggaatgg 15060
aatgatccca taaaaatttt ttcgtcatga ggcaactgcaa tagtgggtag gccaaaaact 15120

gttctagaaa atgaaaactt atcttctctc ttaagtgttt tttgatttaag tttatgaatg 15180
gagaagggtg gagaacaatt tgttttttct gcacaattac caaatcatt agcctgtaaa 15240
tctaagaag ttaaccaaac cgaaccgtct tttaaaaaga tattatctgg actctgaata 15300
aaataactac ctgttttagt attttgatta atatcgtaa cagataaatt gtcacctga 15360
ttatagttaa cataaagaag tccagagccc tcatctagca aaattccatt aggccactt 15420
ccatctgttc catcaacttt attaaaaatta ttgtcactcc aaagtactac atgccctgaa 15480
atacttttaa ataatgaatt cattaacccat tctttcatgc tgatatctct tttatacata 15540
tgagatgcat aaaaacttc atctttttta agggcaacat cattgaaata atattggtca 15600
gggaccctaa tacagccccg ccatatcata tcccatgaag attcattttt tataatttca 15660
aacatttcaa tcgactcaa tggcgaatga ttaattacag ctagtgtata gaagccttca 15720
tcattttcaa caagatctat tccgtgggga ttgaatatgt ctagctcgcc tcttatgcat 15780
gaagcgtcac cccaagagct ttctccaaa gtaatttttg gtaccttttt ttcataaggta 15840
tttaaatcca ttaaggcaaa gtaaccaggg gtatgtttctg catatggacc gatccctcca 15900
aattcagaaa taaagaagaa tttattatca ggcgtaata caatatcttc tgggtttgag 15960

aaattacaaa taaccttgat acggtcatcg gattcacact tactaataatc catttgggggg 16020
cctatatagt cagtagatac gacagttaca gaaataaata aataaagaat agaaaccggg 16080
acgggtaatct tataatagtg ctttataaaa atttctaaaa tctttgaagc atgatttgga 16140
agtgcgatca gccaaaacccc ttttaaaaaa gatagagccc ccataacaac aaaaattact 16200
gaccatagcc tgtcagacca ttccggctga agaagaccag ttaaaaaaag aatcatccca 16260
aagaataaag ctaaatagcc tgagattttg actttcgatc caacaaattt cgaaaagaga 16320
tatgtataaa gaggctttat aagaactagt aaaccacagg ccaagaaaaa gaatgctaag 16380
tagtaattca taagttagtt tttatataaa tgctccttaa taatactaac aagtttctaag 16440
ggcttgcca atggaacatg gtgagcagct ccaggaaccc cctcaaaagt cataatgtca 16500
ccatatgtat ttttaatat ttccaagata cttccggagg ttaataagct gtcttcaccg 16560
tggatgaaca aagcagggca gccaaatgaa aatgtgtaac cgaataacct ttcaagactg 16620
ctaaacatga catcatcaaa ttcccatctc caccagcct caatatTTTT tactgagtgc 16680
tcagcaatgt atcttaagta ccaatcattc gtacaatctt gcttaggcat taacctaac 16740
cttttaataa tatctgtctt gtcttgatag tgcttgatca ttctgagagg agaagagtgt 16800

tgattcgggt cataatccgg tggctctata aatgtatcaa taataataat attatttatt 16860
agatcccttc ttccagatgc aacgtaacca gcaacatgtc cgccgagggg gtgtccaaca 16920
ataaaaaatat ttgaaatat ttttttatcc ttttcctttt caattacaga gacaatacat 16980
tctccaaaat ctttaatgcc atatgaatct ctaaaagaag agtcaccat gccaggaga 17040
tctattgcaa ctatatctgc gcagtctcta aagtggggcg caataggatc ccaccatttt 17100
ttatgagcac ctgttccgtg aataagaatt attaaatctt tgctttcatc tttggagtct 17160
cagctagaat aggatatatc cccatgagga ttcttgataa tctctgagct aggcttgctc 17220
tcaatggcat ctttgaacca ctgtggggca tgaataatgt cttgatttag attgttagtt 17280
atttccataa acagtattct aagctataaa aaataaaaat atgaataaac ttaatttaac 17340
gccagcagca actgttttag tcctaaagga ttctcctgat gggatggaag ttttgatggt 17400
aaaaagatca agtaggcctc ccttcggaga cctttttgtt tccccggcg gcaagattga 17460
cgaaggatga ttcaataata agatagaaga tttttgtgag ggcgtgactg ataaagaggc 17520
ctccataaat ctgggattag attctggagg tctagcatat tgggttgcat gtattagaga 17580
atgctttgag gaggttgga ttttacttgc taaaaaaaag agtggggaag atcttgatct 17640

agatggagtc gataaacata aatatcaaaa atatagagag atgttggttaa ataatagaaat 17700
tgatttatat aaaatctgtt tagaagaaaa tttaattcta atgcctcaac aaatagcccc 17760
tttctcgcac tggataaccc ctgaaataga aactaggaga ttgatacac gtttttttat 17820
tgccccacctc cccaagcadc agaccggaga acatgatggt agtgagctca tagacagtgt 17880
ttggattttca ccaaaagaag cgctcaaaaa atctcgttcg ggtgagatgc ctatgattat 17940
gcctacaata aaaaatttgg aacaatgtgc acaatttgat tcgggctcta agctttttaga 18000
aaatcagagg aatctctcaa atgaggatat cccaccaatc ctgccaaagt ttttaaaga 18060
agatggtgag tggaggggtc tattgcctgg agataaaggg tatgaggatc attaaataat 18120
atggacttaa ttactaaaat aacagctccc aacctgggtg ttttctactgg ggggtgggact 18180
aatacttatt tgattggcaa agatgatata acccttgtcg accctgggtcc aaatatatct 18240
gagcatctag atgaaattat caaagcaggg gatgggaaaa taaaaagaat ctttgtttact 18300
catacccata cagatcattc cccagccgca ttgcctttat caaaaactct taatgtttcca 18360
atgtacggaa ggctagtaga tggatgaatcc tcatgggagg atgaaacatt tatcccagat 18420
attattttta atgataaaga tattattgag acagacgaat atacgttaga agtaatacat 18480

actcctgggc acgcatctaa tcatttatgc tttttaataa aagatacgaa atgcctttcta 18540
acaggcgatc acattatgga cgggtctacg gttgttatg gccaccaga tggcaatatg 18600
acaagctata tcaattcatt agaaaagtta ctagattttg atattgattg ctttgccgcct 18660
gggcatggaa attatatattca tgagcctgag aaaaccattc aatcaattat taggcacaga 18720
ctaacaagag aaagaaaagc tcttagaaaag ctaggagagg caggaatctc atcattagat 18780
aaacttacta agctttgttta tgatgatgtg tcagagatgc tccatcctat agctaaatat 18840
agtctagaag cacatttatt aaagcttata gatgaaaaga aagttaaatt agataaaagat 18900
ctattcgaaa taatttaatc ctttttattt ttatgtaaga ctttctcctc aatagcttct 18960
atatcaatat catcaattga gtcttcgtta ttatcaggta tctttttttac atcttttttca 19020
attttaaggt cgataggaga ggctccaaga tcaaaaagtta gctccccttac atttttttgaa 19080
atagtatcct caacacagtc atcttgatcg tatgcctcac ggggtctctcc tttatcatta 19140
atgggaaagg gtctttgggg agggccatt tgcatgcatt taatagtagc aactggtgag 19200
tagtaaatcgt cactagaata taacttatca agttcttttg gtgatattga gcaaccaata 19260
attccaaata gaataaatgg cgctaagagt cttttcataa actttgtttg ttttctataa 19320

gagactcaac aaccgaagga tcggcaaggg ttgtcgtgtc tcctaaatta gatagatcat 19380
tctcagcaat ctttctaaga attctgcgca taatttttcc cgatcttggt ttaggcaagc 19440
ctggagcatt ttgaattaaa tctgggttttg caatagctcc aatttcttta gcaacaaaatt 19500
gttttcaattc ataactaaag ttgtcatcaa atgatttcatt tatcattataa gtaacaaaaag 19560
catatatcc ttgcccccta attggatgat caaaaccaac aacagctgct tcagcaattt 19620
tagggtgag cacaagagca ctttcaattt cagctgtacc tagtctgtgg ccagaaacat 19680
taagaacatc atcaactctt ccggttatcc agaagtatcc atcctcgtct cgcctggctc 19740
catcaccagt aaagtaaata tctttataca taccaaaata ggtgtcgcgc attcttttgg 19800
gatcaccata aatacttcta atttgactag gccaaagattg ctcaataact aaattacctg 19860
cattagagcc ttctagcgtg tttccatgct cgtcatagag agatgggtta actccgaaga 19920
agggcagagt tgctgaccca ggtttttgttg gagtaatacc cgctattgga gagataagta 19980
cagaaccagt ttcgggtttgc caccaagtat caataacctc gcaattagat ttaccgacaa 20040
cactgtagta ccaatccac gcctctggat taattggctc tccaactgta cctaaaattc 20100
ttaggctatc tctttttgtc tttttaacag gatcatcgcc ttggggccatc agagctctaa 20160

tggcagttgg agctgtataa aaaatactaa tgtcatgctt atcgcatatc tcccaacacc 20220
ttgatgctga agggtagggtt ggcactcctt caaacataag tgttggtgct ccatttgaaa 20280
gaggtcgta caagatataa gtatgtcctg ttatccatcc cacatctgca gtacaccagt 20340
atttgtctc tggccttatt ccaaaaagat attgaaact aatatgagcg cctaatagat 20400
aacctgcagt agtatgtaga acacccttg gcttgccgtg agagcctgat gtatagagaa 20460
tgaaaagagg gtcttcggaa tccataggct ctggagcaca cttattagaa acatctttaa 20520
caagatcttc ataccaaaaca tcttttttat catcccaatt aatttcgcca ccagttcttt 20580
ttataaaccag tgtattttta acatctggac agcccagaag agcctcatct acattagatt 20640
taagtggcac ttttttgccg ccccttaaac cttcatcagc agttataaca attttacaat 20700
cagcatcaag aattctatct ttgagtgatt ctggagaaaa gccaccaaag acaacagagt 20760
gcacggcacc tattcgcgtg caagcaagca ttgcgaacgc agtctcaatg atcataggca 20820
tataaaatac aactctcgag cccttttgaa caccagggtc ttttaaaaa ttagcgaact 20880
tacatacttc gtcatggagc tctttgtagg tcaattcttt agaatacagca gggtcattctc 20940
cttcccatat taatgctatt ttgtttggat cattttttaa atgcctatcg atgcagttta 21000

agctaataatt tgtttttacca ccctcaaacc acttcgcatt attaaattga ttattgaatg 21060
ttgttttgaa gtcttccatc cagcttatgt tttcgtttgc tagattttta aaaaatttag 21120
aaggatcttc tatggattgc ttgtaaagtt ctttgtattc atcaaaagtct ttatatataag 21180
gattacttga atgtttttggg ctataaagct ttcgaggcat tcttaataaa ttgaagggttg 21240
gggggttaatg aaattcttttc ctttgggatt ggacattatt tttgtaatga gcgattcgtg 21300
atcgctatca ttggttttcta aattaatatc tgcagcatta attattaata aaggggcaga 21360
actatagtct aagaaaaacc ttgagtatgc atcattcagt cttccaggt agtcaagagt 21420
tagataattgt tcgttaatat ttccctctctt agtaatcctg tcttttaaca catcaatagg 21480
tgcctgaaga tagattacta ggtcgggtgt tggcgcgtcc agggtagat ggtcatatac 21540
tttgtcatat agatccattt cctcgttaga aagagtaacc tcagcgaata atcgatcttt 21600
ttctattaaa aaatcagcaa ccctcactgt ttcaaaaagg cttctttgtt taagatcttg 21660
aatttgttgc attcttttgaa acaagaagaa aagctgagtg gctagagctg atgggcttg 21720
gtttttataa aaattcttta agaaggatt ctctgccggt tgttctaaaa aagaatcata 21780
attaaaatgt tcggctatct tatttgctaa agtagtttc ccaacaccta tcggctcctc 21840

aattgcgata tatttttgga gtggtacttc tttaatagct gggttcattt ataacgtata 21900
gtttttatcc tagattatct caactcgtta tagaagcaa atatttttgc agattcttca 21960
tcaacttttt cattcatctc agcaataatg gttcctctca atgagtcatc attttttttgg 22020
aaatcagtcc accacttccc aataccactc ttgtctccag aagctctctc tcttacaagg 22080
agaaaaatcat atgcagctct aaaccttgga tgtcgaagag ttttatatgg ctggctacca 22140
attctgctat gaagttttaa ttgaagtacc caaatatcct taatatagct tgaaaaatttt 22200
cttggtattg ctgtgatatt ttgtgttg cgaaagtacac catccataga tcggaataat 22260
tttctcacat taatttctcc attcttagaa cactttttca ataacgaggg ccataacaat 22320
gcagccataa gaaagcctgg tgtaattgac tggttggttt taaccctatc atcagtattt 22380
cttagagcat gtgtcataac attgctggca aagtcattcc tacttggatc agaaaggata 22440
agatatattat ttaaatgaaa tgagcagagt ttttcaaaat tcttttcacc cattccattc 22500
aaaaatatatt tgcagaactc atcaaacaaat ctagcgtttg agatgcctga taaaagatgg 22560
cctttatcat agatggcatc cttaaccaga ttatctatct taaaattgag tttgttacta 22620
aatcttatag ctcttaagct tctcactgga tcttcttcga atcgtctttg gggatctcca 22680

atagatacaa taacctttctt gtgtatatgc ttaagcccat cgttatgac ttctattttt 22740
tttghtaacag gacagtagta aaggcatc acagtaaaat ctcttctatg acaatcttgt 22800
tcaagagtgc cccaattatt atctctaaga atctttcctg tcgaatcagt aacaataatt 22860
tcaccatctt cctgactact tcctgatctg aaagttgcta cttcaagcaa ttcactccta 22920
ttaaaaacat gaaccaattt aaatcttttg ccaattattc ttgaagcctt aaatgttttc 22980
ctgatctgct caggcgtggc attagtggct atatcgaaat cttttggctc tagtcccgtt 23040
agtgcacgc gaacgcacc accaaccaga taggcctgaa aattattttt ttgaaggctc 23100
tgaacgacag atatagcaaa ttacttatac ttattattat ctatcaaatc taattatgaa 23160
tttttatgat ttaagggcat tagccattta gctgcttttc tttaatttcg tcaagtgtt 23220
tgcaatgaat gcaatgagtt gcagtggtc tagcttcaag cctcttaatt ccaatttcat 23280
caccacagga ttcgcaccaa ccataatcat cttgcttaat ttgttcaata gatagacct 23340
tttttgctgat aagtttttctc tctctatctc tggttcttaa ttcaaaagca aactcttctc 23400
cttgagaagc cctgtctact gggctctgcat aggttttcacc tttagctcta agatgatcaa 23460
aagttttttg catttcatcc ttaagatgtt ctttccaaag aagaagaacg gcaacaaaat 23520

gtttcttcat tgctgcactc atatatTTTT cacccttctt agatttataa ggtgcaattt 23580
tagatttatt attagctatt gttgcttttg cagacttctt ggctgcaact tttttaactg 23640
gagctttttt ggtaacagtt tttttaactg tagctttttt ggcaacagtt tttttaactg 23700
gagctttttt tactgtcttg gatttttttt cgaccatgta agattttata gaatttttag 23760
gtggagaaaa taccagatac tgacaaaaatt agctagtcat ttttaattt atttaatact 23820
ttcaagtacc catccgagct aagtcttggc ccaaatgtct caacaacctt ggaagatgca 23880
tagtttgcaa acttagcaca tgcttcaata ttattccctt gaaggtaggc atgcataaac 23940
gatccggcaa acatatcacc agctccattg gtatctattg gagttatttc ttctgcttga 24000
gcatgcttct caacccttt atctataaca acacttccat cggcgccttt tgtaatagca 24060
gtcatatagg gctttttcttt ataaaagcta acagcatcat caaggcttct tttaccagaa 24120
aaagcaacag ctcatcatc attacagaag atcatatcta ttccatatga ctctattaaa 24180
tcaaatTTTT cttaaaacc atgaacaata cctgcatcag aaagagacaa ggcttttcttt 24240
acgtccttgt cttaagggtg ctctaagact gaaataacag cattaaaagt atcgtcactt 24300
gttaccatgt agccttcgat ataaaaaatt ttngaattt ctacaacatc aaaatctata 24360

tctgatttac caagatacgc actaactcca agcatgctgc tcatagtttct cttagcgtca 24420
ggagtaacta aaattaagca ttcccagtt ggttgatctg tatttcaga gctgacacca 24480
atatgtttga ctccagccga cctgagacta tcaagatagt ttcttccatc ttcatcatca 24540
gaaactctgc atacatgatg gcaattcgaa ccataatttg ctgcagcaac aagagaattg 24600
gttgcagagc caccgcaatc agaaatcgat tcggctccca ttccaataag tttgctaatt 24660
atagggtgct gttcttcaga agatgaaaga gtcataagat cggctacaag gcctacactt 24720
gataaaaatt catggctgac ttatatatga gtatctacta aagcatttcc aaggcgcta 24780
atatcatatt tcatgtgtta ttcccttttg tattatttg ttactgttt ctaacgttct 24840
attaatttct tgatccttat gcattgctga aatgaagcct gcttcattatt tagagggggc 24900
aaagtatat ccaacttctaa tacatgaatt taaaaaattt gaaaatagta catcatcagt 24960
ttttgcaaca tcattaatat tattaggag ttcttctgaa aaaaagaatc cgaacattcc 25020
accaattctg tttatagaaa acggaatacc tgactcaatc attaagggttt tcatcccatc 25080
caaaagcaca gatgcatttc tttctaactc tttaaacgga ttttctttaa tcaacaattg 25140
caacaaagca gttcctccag ccatagctag tggattgcca gacaaagttc ctgcttgata 25200

gacaggacca gaaggagcta gatagttcat aatttcttct ttgcctccaa aagctccaaac 25260
aggaagaccg ccacctatta ctttccccaa agcagttaag tcaggagtaa tattataaat 25320
ttcttgagct ccgcctagcg aaactctaaa gccgctcatg acctcatcaa atattaaaaat 25380
agaattattg gctgagggtgg ttccctctcaa taactttaag aaatcttcat gacctggaac 25440
aaagcccata ttccctgcta ctgggtcgac tattacggcc gctaaagtcac cttttatctc 25500
attaataatt tctaaaaact gttctttatt attgtattcg caactaaatg tataatttgc 25560
caaatctgca ggaacccctg gagagtcagg taagccaaag gtggcaaccc cagaacccgc 25620
cttaattaaa agagagtcaa catgaccgtg ataacagcca tcaaatttaa taattttatt 25680
ccttcctgta aaacctctgg ccaatctaata cgttgtcatg gttgcttcag tacctgaatt 25740
aaccattctt attttttcaa ttgaaggaaat gcatttttta attagcctgg ccacatcaga 25800
ttcaagactt gtcggggcgc cataactagt tccaagcgca acttgatttt taattgcacc 25860
tacaatgtct ggggtgtgaat gaccattat catcggaacc caagacccaa tataatcaat 25920
atattcatta tgatcagcat cataaaggta tgcgccggat gctctttcaa aaaatatagg 25980
attgccattg atatttttga atgcccttac tgggtgaattt acccctccag gcataagagt 26040

cttggcctct ttgaataagg ctatggattt atcaatttta ttatgtgtca aagctaattc 26100
ttcttaattt ttaatttcga tatgatatca acttaattct tattatgtat ttaattgttt 26160
tgaaaaattca tccacagtat tccagttggt gaactcataa gtattttgtaa catctgtagg 26220
accctttgta atccacataa tcaatctaata catgtgttta tctataaaatt tatatttttg 26280
ataatctatt ttgcctgcaa agactgctaa tttcattggg ttccaaggag acagctctaa 26340
aaatttttgc atatacgggt ttgttttcagg tgtatttttt tcaggccttc tcgcaactac 26400
attaactgaa aagaaggcat tttcttttgt ttcaagacaa gcaacatttt tttgaataaa 26460
ttcataaagt tctggtttgt gttgccata cctaagtctc gcaccaataa taattttatc 26520
aaattgatat aaatctaact ctactgcttt tgcaatatgt attatttttg aagattcaga 26580
cacatctaata ttagagaaaa tttttttaca aatttccaga gtttgcccat cggttgtaga 26640
gtagattagt agagttgatt tcataaatat attttatatt gttaatttaa aaatttcatg 26700
gcacgcgcaa aatataaaca aaaatatagt atatttaaaa aatgattaaa tatttattta 26760
ttatcgcttt attcatcagc aactttgcat atgcaaatat tgatgcagct aggtgcgcag 26820
ggattagttc agatcaagaa agacttgatt gctatgattt aatttttaa gctaatgatg 26880

agctaccttt agatagtaat attaaaactc tcatacacacc agctatcaag gctgtgacccc 26940
cagctgattc aataaaaaatt gaaaataaag caacaaaaga aaaggatttc gggcttccaa 27000
aaacaaaaat caaaaactct gcaaagaact caataaaaac ctcggtagta aggattaaaa 27060
aaacaaaaag tggtaaatat atttttactt tagaaaaatga gcaagaatgg actgctgaaa 27120
cttcttatag agcaaggaat atgtttaaac cagaaaccgc agtcatttta gaagaggccc 27180
tggtttagtg tttttatatg attaatataa gtaataaaca gaaaattaga ataaagaggt 27240
tgaaatagca atgaccatac agagtatatg aacgaccgag ggtgccatga aaagaataag 27300
atctgttctg gggtcccaag atggctcaag cttcagggtt tatgttaccc gaggtgggtg 27360
ctctgggttt caatatggct tcaagtttga taacgatatt gcattcgcgt atgatgttat 27420
aaattgtgga gatttctcgc ttttaataga ctctatgtcg taccctatc tttatggatc 27480
aactctggat tttgttgaa atctctcagg ggctaaatc gttattaaaa acccaaatgc 27540
caaaacaaca tgcgggtgtg gagagtcatt tacagtttag attttgtaat tgaacccaaa 27600
agtcctttta cattcttgct tgatgtctga actatcagag cctcattact gactcttta 27660
tatcccatcc aggcaaatgc catggactca atagcaaaa catcatggcc taaatcactt 27720

gagaggacga tatcattact agccatctca gaaattcttt ttaccagata cttgttatgg 27780
gcaccaccgc cacaaatgac aatgtcacaa ttattatgac catttttatg gattgaattt 27840
attatagatt ttgctgaaaa ttcaaccagg gtacatagaa tatcttcagc tttctttttt 27900
aagaatttct tggataatat ttttatatta aagagctctt tgcctgttga ttttgacat 27960
tttcttttaa aaaaattatt ttgaagcaat ctccctaact caatgtgac tacttttct 28020
ttggctgcaa ttgcaccatt tttatcaaaa ggaatctgaa gaaatcact acaatatgca 28080
tctaaaaatg cattaccagg gcctacatca gttccccata tgtcatttct attttttaca 28140
aatgagtaat ttgaaatccc ccctatattt aaaataatc gtgggtttct agctttataa 28200
aataattgat tatggaattc cggaacaaga gggggcctt cgccaccag agcaatatgc 28260
atatttctaa aatcacttac caccagaagg cccgtttctt ttgcaacaat atttgatca 28320
ccaattgca tagaaaatgg aaatctctta ttaatttcat gcctaattgt ttgtccagaa 28380
atagcaacac actctataga tgattttctta attttactaa atccaatcgc ctcatataa 28440
gattttgaaa ataaagaatcc tatctcttta ttgatatgac ctaaacttga caacgagctt 28500
tcgttatttt caataagttt ttttactttt aatcttaggg attttgggaa ttaattgaa 28560

tggaataact caagatgtat ttttgtacca atgctaagaa atgaaatatac aatagcgtca 28620
tgactagtgc ctgtcattgc tccaatatag atcttttttac tcattactac ttggagaaaag 28680
tttattgaac tcctccattg ataatttatt attctttaat agtgatatcaa aacttgcacct 28740
taaatTTTTT ggtactgggt cagctgatgg taatttttact tttataggat ctattcttttt 28800
cttacctatc ttaaaactcat aatgtaaaatg aggccctgtt gccagtcctg aactcccaac 28860
gaaaccaata gtatcgccct gagaaacttt tttgcctttt ctatatccct tactaaaattt 28920
ttctaaatga caatatctag tcgaatactc atttgtatga tttataacta tctcattacc 28980
acatccattc ctttgaccag aaaaagaaac aatcccatcg cctgtagttc tcaccggcga 29040
tcctcttttt gctgcataat cgactccatt gtgagctctt attgtatgaa gaacaggatg 29100
cattctgttt ggattaaaat gagaaactaat atatgcaaaa tctaaagggg ctcttaaaaa 29160
ggccttttgc atattatttc cattttcatc aaagtactgt tttttactgg cttcagtga 29220
aaacctattt gcaaaaatagg tattaccatt gttaataaac ttagcaatta cgatatcgcc 29280
atttttttact ttctctccat cactataagg agtgtcataa attacatgaa attcgtcacc 29340
ctctctaata tcaaaaaacga agtctacatc ccagccaaag atataagcaa agtccataat 29400

cacactttcg ggaatatattg catctagagc tgcctcataa aatgaggatt gaataatccc 29460
gctattgtaa gattcaatta actcaatact tttgcttata ttcttaaata ctatttctgg 29520
agtaaggcta attgaaatag aatttatttg atctttcatg atctcaattc tggtagctc 29580
ctcaccagaa tattcaaaga gcattttttc ccctggttta atgttagcaa tgatatTTTT 29640
ggagtctaata ctaaaaattt tataagcagt atttaaaggc actgaaaagt tttcaaaaat 29700
tattgaaaga tttttctccat cttgcacctc atgcattctga tatgttttctg tttgctcgag 29760
aggtatagag aatttttccag taatctcaat ttcttcaact ggcaaaagatt catatgttcc 29820
tatatcaata taaagcatta atattaagac aatagaaaac gcaaaagaaa caagtaccgc 29880
tctttttgga acttttttaa aacctatcat tagtcctctt gaatatTTaa gagttcagtg 29940
ttgccaccac atgctaccga gttcttttgag acaacctttt cgTTaagaaa ttgcagaagg 30000
taattaggcc caccagcctt ataaccgctg cggatagat tttgtccacc aaaaggctga 30060
gaaccaacaa cagctccccc catatctcta ttatgtaaa tatttccaac attacacttg 30120
tcactaaaata tgcgggccct tttttctact ctagtatgaa taccatttgt aaggccgaag 30180
ccactatcat taatatTTtc aattagtgc tccatttcat ttgattTaaa tctaacgata 30240

tgaagtattg gtccaaattg ctcgctcttt agatcagaaa tattatcaat ctcaattatt 30300
gttgagcta caaaattttc atctacatta tcagcagacc taaaaataga ataattctta 30360
tcttcaaac ctttgacgta agcattgaga ctatcaagag atgttttgct aattattgga 30420
ccgatatctg tgtcaagatt ttgagggttt ccaattttta gttctttcat accaccctta 30480
atcattgata aaagatcgtc atatatctcg tcttgtagac agagaactct caaggcagag 30540
catcgctgtc cagagctatc aatgctgaa cgaataatat catcgggttg ttgctcgaga 30600
agtgcgcttg aatcaacaat cattgaattt atgcctcctg tctctgcaat caaaggaata 30660
attgattcat ggttacttgc aaggctgctt tgaatttttt ttgctgtttt taaagacccg 30720
gtgaatgcaa cccctttgat gttattaacc ttgaaagca tgtcaccatg aatgccgtct 30780
cctaggatta aatttaaagc atcttttggg accccaatct catgaaaattt gttcacaatg 30840
atgtagccaa gaattgaagt atgttcttgag ggcttaactg ttactttatt tccacatgca 30900
agggctgcac ttatttgccc tataagtatt gcaacaggaa agttccatgg actaatacat 30960
aaaatatggc cttttggtga gtaagatagg gcatatgatct cgccggtggg accctcaaga 31020
atatggtctt ctgtttggag gccaacagct tgttttgcat agtatctgag aaaatctata 31080

gcttctctaa tttcatcaat agtattttgt actgtttttc cggcttcatt cataaggtaa 31140
taaatagct cagaaggatt tgcttcaatg tcgtctgcaa ttttttctaa aatggatgct 31200
cttttttcaa catgcatcaa agaccattca ctaatatattt gttcttcaag gctgcccttt 31260
atataatcta gatcatcata tgaggctgta ccaatatattc ttccatctgc aagagatgag 31320
atgtcatgag tattgggttt cttataatcc tttgctttat aaatcgatga agcctttatt 31380
tcttttgaat caaactttcc aagctcctct tcaagcattt ctaaattcac cttttcactt 31440
aaatcaaatc ctttagagtt cggctctgtca ttaaatatat ttcttgcat tggatatctt 31500
tttttttcat cctcaatttt taaatggggc cctctggcaa gccaaagcaga atctgtttca 31560
ggatcaagta acctattaat aaatgagctg tttgctccat tttctaatag tcttcttaca 31620
agatagggt gcaagtcttt atatttgcca atcggcgcat aaattgaggt attttttttca 31680
gtatttaaga ttttgtttgc ggacttataa agcagctctc ccattccaaa tagtctttga 31740
aattcataat ccttatttga gccaaagatga tggattgcag aaatgggatg tgcattatgt 31800
gtagcaaat tagggtaaat tttttccaca ttaaaaaatc tttttgcaca ggctaaataa 31860
gctaagtctg taacagattt ttttgtag acaggataac catcatagcc atagatttga 31920

gc atg ct taa ttt cataaat cca at atg ct ctt ttacga gcc gaacatg cat aggtgct 31980
ctgttttcta gtagctcttc taaccaatct attgtggcta tagctctttt gccat atgcc 32040
tgaacagcaa taccaaatcc tttccagttt ttaatatattg gtgaaagggc catctcttta 32100
ataatctcta aactaacggc cagtctgtct tgctcctcag catcgattgt aattttctaca 32160
tcttttagact ttgcatactc tgtgagctga ataaagtttg gaagtagatc agacttttatg 32220
tcttttagct ttttcatttc atatctagga gataaagctg atatttttat tgagacacca 32280
ttaattgtat tttttgtcag atttatcttg ccgacttcac ctatagcatt cttataagac 32340
tgataataag tgtcgggcttg ttcagcgttt ctctgctgcct ctccaagcat gtcaaatgaa 32400
taaatttcat tttctatatt cttaattttt tttatgtcat caaaatctct gcccataaca 32460
aattcttgac tgagaatatg catggcacca actacagcat ttctaattgg aaactcgcca 32520
gattttgaga ttaaagaact taataggcg ctggggtttt tgggtccactc atcgggtgta 32580
gaaactacct tgcctgcaag aagaagaccc catgttgatg cattaacgaa aacactatcc 32640
gctttgttta aatgctctat ccacgcacc tcagataact tctcagatat tattagatcc 32700
cttggtttttt tgtctgggat tcttagtatt gattcggcta aacacatcag agcaaacacct 32760

tctttattat ctagcccata ttcgcttaga aatgcata gttttgttct ctgcgtctta 32820
ttctctctgc aagcatcaat tattttattg agcatttttt gaaatcgaag ggtcatttaa 32880
aaaatcagaa ttactaataa gatctgaaac tatttcttgt tcaggataaa acttgttact 32940
tgttaatgtc atagctatta ttttaactct atagtgccaa gtatcaactt actgatttcta 33000
ttaatgactt aaaaacatta tgatagtcac atgagtgatg cacttaaatt aattaaaacga 33060
ggaaccgacg agatcctcac agagtctgat ttaaaaaaga aattagattc tggaanaacag 33120
ctaatcatta aggcggggtt tgatccaaca gctccagacc ttcatttggg tcataccgta 33180
ttattaaata agctaagaca ttttcaagat cttgggcata aagtaatttt tctaattggg 33240
gattttactg gtcaaattgg agatccctcc ggtaaaaaata aaactaggcc aacacttact 33300
tcagaaagaat taatttcaaa tgcaaaaaaca tatgagaaac aagttttttaa aattctttaa 33360
aaagaattaa cagaagttaa atttaattct gagtgggtgca acaagcttgg tgcagatggg 33420
ttgattgggc ttgcatcaaa atataatgtt gcaagaatgc tggagcgtga tgatttttaat 33480
aaacgttata gcgcaaatca aagcatagct attcatgaat tttataccc ccttgttcaa 33540
ggatatgact ctgtagccct agaggctgat gtcgaatgcg gaggaacaga tcaaaaaatt 33600

aatttgctag tagggagaga gctacaaaga tcctatggtc aagaacctca ggttggttta 33660
actgtaccca ttctagaagg cttggacgga ataaataaaa tgtccaaatc attaaataac 33720
tttatagcaa tagatgaaga gcctaatagat atgttcggta aaataatgtc tatttcagac 33780
gagttaatgt ggagatggtt tgagttactc agctttacct cagagaaaga aataggagtt 33840
cttaagaaga aaatggaaga agggaccaac ccaagagata ttaagtttct ttagcagaa 33900
gagtagtag atagattcca ttcagagggg gatggttcga aatgcaaggga agcttttctt 33960
caaagattc aaaaagggtca aatgcctgat gacattcctt ccatgtcagt tgatgttggg 34020
gctgagggca ttccattagt aaacctttta aagaattgtg agatgacatc aagtacatct 34080
gaggcgatga gacttggttaa acaaggcgga gttaaaaatcg actctgtaaa aatagaagat 34140
cctaaaatgc taatttcaaaa aggccaggag tctatttatac aggttggcaa aagaaaaattt 34200
ttaaaaaatta aaacataatg aaaaataaat tagtccaaat atttcttctt ctttttatag 34260
ttgcggtgtaa tcaggatagc cctaataatta aacaaattag taatatgcag tattttattg 34320
ataatgaaat aagggaagga atctattctg ttgaaccagg ctgcaatac tcaattatcc 34380
aaaatggaga tcaaagtctt gaatcaccat tgctgcagga tacaattaca gctcattttc 34440

acgggaccct cactgatggc tcagtttttt ggagttctgt tgaaatgggt gagcccttaa 34500
cagtcgaact atcaggtcta atagttgggt gccaaaaaat aatctctatg atgaagaaag 34560
gtgatgaatg gagagtttat atcgacccaa gatatggccta tggcgatgag ggccgggcctg 34620
ggataccttc aaactcaatt cttattttttg atattgagtt attagatata caaaaaaact 34680
aacccttata tataacggat agagcgtatc cataaacctc agcaacctgg ttaataattt 34740
tatttttagg tgttcgggct ccatgaccag ctcttccttc aattctaatt aatattggat 34800
tatcacaccc ctgggactct tgaagttttg ctgcaaaactt gaatgaatga gagggacta 34860
ctctatcatc tcttttagct gtagtaatca gagtagttgg atagcactca ccttcaacaa 34920
tattatgcag aggcgaataa gctaataagt tttcaaaactc atcttttctta tctggagagc 34980
catagtcact ttcccaggcc cagcctatag taaatttatg aaatctaagc atgtccaaga 35040
cgccaaacttg tggaaattgcg actttaaata aattaggatt ttgcaacatg gtcgcagcca 35100
ctagcaaaac tcattagag ccaccttgaa tagcgggtga tgatggagaa ccaatttctt 35160
gtgcgtgtaa aaatttagca gaatatgcaa agtcatcaaa aacattttgt ttattaaaaa 35220
gcctgcctgc atcatgccag ttatcaccat attcaccacc acctcttaga ttaactacag 35280

caactattcc cccttggttc atccaagtga ggtagctttt actaaagcca ggtagtcttg 35340
agataattaa gccaccatat ccataataaa gtattggagt attactgtct atttttaagg 35400
actttttgtg actaagatgg atgggtattt gtgtttccatc tttggatgga aagaatttaa 35460
agtcagatgt aaataaaagtt gagtcgtgtc ctttaagatc ttctttccaa aaaagtctctt 35520
gtgacataac agttagatta attttgtata tttctcgagg agttacaaaa ttgtataaatg 35580
aaaaataaga tacctcatct tcaatttcac caccaaaacc acccattggt cctttttcttt 35640
ctgttgctag cttattttta tatgctcctt ttaaatcaaa gaagtgaacc tcagtaaaag 35700
tatcaacgag ataagaaact acaatagaat tatttataaa gctaacacta ctaatagaat 35760
tagtactttc accaagcact tcattccaaa caaatgagcc attttttatt gtaagtata 35820
ctacttttcc gttcgcagca ttttcagttg agtaaaacca aaaagtatca ttcttgcttt 35880
ctaaaaagct ataagcgctt attagctcat ctattaaagg aataaaaagg tgatcagggc 35940
ttagctgaac atagagtcta tttcttttcat ctgtgccttc accaatagat agaaacttaa 36000
ttttagaatc ttttacaaca cttattcccc aactccatct tggcttttca ggattctcat 36060
aaacaatcac atcctcatct tgagcagtag caattttatg aaacattaac ttggagctg 36120

tattaatatc ttttaaaagc tcttcgatg gctcgtcgtgta tttttggtaa taaaatccag 36180
aatcatcatt ttcccatgaa gcaccagaaa atttagccca ttcaatccta tcatcaagag 36240
tttttctctga ttcaatatca agtactttcc aggtttctcca atcagaccca ccattctgata 36300
ttgagaaaagc tagaagagat gcattcattac taacgctggt acttgcaaga gagatagtctc 36360
catcttctga gaactgattt ggatcaagta aaactctatc ttggcattct tcacagtcct 36420
taatcatcag ctgcttttgt tgccatgacc catcattgaa ataataaaaa gtttttttgt 36480
taacctgata aggcatgctt atcgaatctg tatcccaaac ctcatctaaa ttcttagcaa 36540
tagatttttt atatttattt tggcctataa atttttttgt gaaattatatt tgtctctcga 36600
cccagtctgt tgagtcctca cttgtaaaaat cctccatcca tctataagca tcttcgatta 36660
ggtatccatg gacttcttca ctaaaaggaa cttttatttga ttcagggtat tcaaaaattat 36720
ctgttttgatt tgagcaacta actagaatta atagagctgc catgtatatg tattttattgc 36780
gcataaaaag cataatactt tacaattacc ttatgaatca attcaatcga gcatggcaac 36840
ttttacgcaa aaaattggaa gttttttatt gtacttgccg caccatcat gatgcttgaa 36900
atagcaatgg catcaatggt aacccaata caaaacgtta ctcagcctga ggatattcta 36960

gaatttttca atgagaatat agccttttcta ggctctgtta gtctcttagg cgtagtttta 37020
agtatggctt ttatgggagc gctttttgtt tcgtatgcat ccatagaatc agaaaatgag 37080
attgagcctc taaatgcctt atttttaggc ataagaaaaat tctttccact tctggggggct 37140
tatcttatag cttcagttgg tgtctttttt ggtatcttat tattaatact ccagactttt 37200
tatgtagcag caaggctttg ctttttcct gcattttatta tgcttgaaga caaaggagct 37260
atagaatctc ttaaatatc ttgggaaaag acggatgagc atggcaccac tttgtttggg 37320
cttaccatta cttcttttc tttaacaatg attttgcac cagttgccc atccattata 37380
agtccgggat taatgcaatt agttgttctt gcaattattg aatatgtaat agtaattcca 37440
tgggggctatg tataatttag ttatacaag tcattaaaaa gatattaggc agattaacta 37500
gaataaatta atacgaagct ttctgtttac aaaaaaacac accactctta taatccttcc 37560
tccgaagatt tgcactttgt taattcatgc aaacaacttt gggctctgtag ctacgcttgg 37620
ttagagcgca cccctgataa gggtgaggtc ggtggttcga gtccaccacg acccaccatt 37680
ctttgaacat taacttttat ttcctttata ctattctta actaaataag gatagtatta 37740
tggctaagg ttagacaag caaaaaaacg acaagaagaa aggcaaaact ctaaaaaaa 37800

aaagagcggc aaaaaaagaa aaaaagaagt agttttttgt agaagactat tattatctag 37860
tcgacctagg gccatcttta aaatctatctt taaagaacaa cattattccg gtcactgagc 37920
ttactaaagc taaaatagaa aatattttta aaagaagggt gtcgatatata tctctttctt 37980
tccaatccat aatatgaac cccaatca gatcccatat ttccattta ttagatctaa 38040
tggccacaac ctctgctgag taaacattaa tgtaaacatt taattcagta cctttaacat 38100
ttttactttt tactctataa attggcaatg atctgcctct atattcagat cctgattttt 38160
catttacaac ttcttcggta gcaaatggca aaagagtgtt ttggcttgaa actgaatcta 38220
tagcatcttc catagaaatc ttgtttaag gctgccaag catattcaga tacttagttg 38280
agccctttgt tgtgatgata acaatctctt gacctaatct ttttttgaat ttaacttctt 38340
tagctttctc tatttcaaaa tttaacttac ttaaatcaaa agatgtttca acatgattaa 38400
ggtattgctc gcctcttacg agctcaatct tattaaaagc aaagtatatg cctgagatag 38460
tccaaaagtaa taactgaaga gagataaaaa aacttaggta cttatgaatt ttcctaacta 38520
aaaagttcat cttttcataa tcttcattat ttcataaatt ttacatcaa attcgtcatc 38580
acctttttaa gattcccttaa tgcattcacg caaatgggtt tttagaattt tgccctcaac 38640

agtaataatt gaattcttaa gtgcttttat ttgattaaga atatcaagc agtattgacc 38700
ttcctcaacc attccttgaa cgcctgcac ctggccctct attcttttta agcttattaa 38760
ttgctcttta tgacatggat gtttcatttt ataaaaacct cggaactata tacgtaaaaa 38820
ttgtagcgaa taaaaaatt gctacagatg catagtacag atacctagat ttttttctag 38880
tttctgaaca tgctctgcct agttctgggt cagcagggca aggcataatga taagttttgt 38940
agttgatata tcctgcaatg acaatcatta ctaatgcaaa tagtgttatg taaagcttat 39000
attgagatag ggtaattaa aatggaaata cagttaccag gcttgcaaaa ctgccccag 39060
caccaagagc aacaaaaata gctggcaagg cacaaatat caaagtagaa gaagatgcaa 39120
ataatgaaaa aaaattcgac gccttatcat tcatagttat taaatttcta ggatttgaag 39180
atcagacata ttttggccat tagaaaagaat cttcttttcg atctcagaga aggtgatttc 39240
ttttgccta gaaaaggcaa tcactacttt tccattatca agatcaacat ctattctttt 39300
gacatcttta tcttttagaa aagttttttga aatacctctt gcacagaaat cacaaaccat 39360
tcccttcacg cttacgatcg ctatatttac atttgagagt ccttcaacaa atgcatcaaa 39420
cctctcaggg ttacatcaa gatcttttcc atcaaccatg gtttcatgta aatgaccttc 39480

atgagaaatgc ccatgcatac cactatggtc attctccgcc gcaataaggc cgcttgaaag 39540
caaagtaatt aataataatt tttcatatt tttctccttt taaatctgt acataaaatg 39600
taaatcccag ctgatatctg cgtcatagcc aaactcaaca agggcattac cataaaaaa 39660
tttcaattct ggatacgtag atgatttatt tgttacagta ttttttttgg ttttaaccat 39720
aaccatgta tgtaaatcat tgtagtcagc aacgtaagga gcaaatccaa gttgaatata 39780
ttctttcttta tagtctttgc taaatttact gttaatatcc tttaacccaa agcccgcata 39840
ccacttccgg gtttcccagt ctccatggat accataaaaa ttatttttaa tacttccagg 39900
cacaattcct gattgaaaat ataaattcct ttgggatgta aaagtatttt ttctattcac 39960
caaataagta aaacgaaaat agttttctgt gttcttataa aatttatttt tagccctttc 40020
taggccaatg gaatatattgt aggttggaga gtagtggaata taaatagaat cattaaaggt 40080
atctgattta tacattatgg ttgttccgcc agagtaggat atcgggtcttg catcagcatt 40140
aaagaccatt aaaaaaacta aaatgattat tttatatata cccatggggg gtatcttacc 40200
ataggattat aaattaattt caggcatctc tccaagcgcc cattttcaat gaaatcacca 40260
tcagaattcc aaaaataaacc atagctatag cagatgcata aaaaatacta actagcccat 40320

atgagagcca gtaaacagct accaatgcta taacagaaaa tactataagc ctgatggccg 40380
tgattataat tggccatttc attgcattgg ctccctgaga tgcaaaaatac agggaaagcc 40440
catatccttg aaatacataa catacaccaa ggatttgaat atactgcttt gtaactaata 40500
aggtctctgg atccgatgta aaaatactaa tccaaagatt tgggtgttagt gctaaaagcta 40560
gaccgataac agccgaaagc aatccagctg ttgttgcacc aaccatgcct attttttctg 40620
ctcgcctctat attttttggca cctatatattg ttccaaccat ggctgtcata gcagtcccaa 40680
tgccgaacac aattgggatt aataagaact ctaccttga accaattcca taaccggcaa 40740
tagcagacgt tccaaattga ccaatgagtc ctgtcagtaa aagaactgtt gcaactgtca 40800
ttattggcga aagagatgcg ggtaaggcaa cagaaaaaat atcttcgaat aattctcttt 40860
caattgttag tcttttgagc ttaagcctta caggcgatga agggctcgaa atttttatta 40920
aagttactaa tgccataaaa cctgaagtaa ctatcattga ccaagcagat cccacaagcc 40980
caagtttagg tagaccgaat gatcctaata taaatccagc agagaaaaaa acttgtatac 41040
ctgcacaaat taccgttagg actgcaggaa attgcataac tcccatccct cttagggcgg 41100
cggttaagct acctgatagc caaacaacta ttgctcccaa taaatagaca aaacaataag 41160

ctaaagactc ctctaataat gcacctgttc cacctaatat tcttaataga ggctctctcaa 41220
aaataaaaa gataattaaa aaagctaaag ctccaaagca agatatatac aaagaatgcc 41280
ataataattt ttcagcccta ggcttgtctg cggctccaag acttcttgca atagaagatg 41340
tgaccgctcc tccaagagca ccaaatgcca tctgctgagt cagcatgatt gcaggaaatg 41400
ccaaagtaac agcagctaataa ggagttatgc caagttgact aataaaccag aattcagcaa 41460
gaacaaccac agcattttatt aaaaatgcaa cagtattagg tgctgacatt ttaatcagca 41520
aaggaaatat tggatctttt aaaaattgct ctgttctctt gtccataacc ttctcattta 41580
ttaataattt tttgggttaa gggtgaaaga gaacacgttc tcaggattca cgaatcttta 41640
aataataata acagctttta gttaatttaa tttttaaaaa caaccaaatc tattatctgc 41700
tagattttat tttttaaaca tattaatatc taattttaatt gcaaaggggt ggccatttcg 41760
gcctgagatc cttggaaact ctttgaacct gatccataca ataatggcg aggaattgca 41820
tgaaaaatcat taaaaatcat ttattttatt caataatctt actatatatt tttagtctag 41880
agacgagctc tcaagctata gaagaagtta ttatcaaagg agactggaga gaaactagtc 41940
tgtcagcaga agactcaagc attgcagtcc ttgatagtaa attaatagaa tcccaggctt 42000

taaagcactt tgaaaacctt tcgtacctag taccaaatTT aaattttgct gcaagtgatt 42060
ctagagcaag acatTTccag ataagaggaa ttggagaaag atctggctat gaaagaaactc 42120
caaaactctgc agtaggttta ctgattgatg atatagactt ttctgggcaa ggtggaattg 42180
ccacaacttt tgatgtttgat caaattgaag ttcacagagg cccccaagga gcaagaatag 42240
ggtctagcgc aatggcgggg ctcatatata tttctacaaa agacccaaca gaaagctttg 42300
aagggaaaggg tgaaatagtt atgggggtcat atggaacttt taataccgga attgctgttg 42360
gcggtcctgt aaattttaat aaagacctca cttataggtt agctataaag aaagattatt 42420
ctgacggatt taggaagaat atttttctca acaagtctga tacttctaag aaagatgaga 42480
gcacatttag attaaaagtg aattgggtaa ctgataacca aaccacttat aagtttctta 42540
tatctcaaat agagttagat gatcctgcag atatatggac tatagatggg agtctttaata 42600
cattatctga taagcctgga atggactctc aaaaaagtaa tgcctatggg gtaaaaaattt 42660
atcatcaatt taaaaaattt gaatttcaaa gcctatcaag cttaacgaat actgatgtca 42720
ttcttagtta tgatgtgat tggggcaacc ctgagtcgca ttcacccttt atttatgact 42780
atttttcaga aactacaaga aaaagagata ctttagtca agaatttaga cttgtatccc 42840

aatttgcaga taaaaataca gaaaaaagca tcgaatgggt cgttggggct agttttgtag 42900
atataaatga aacaaatgct aaaaaagata ctggtattta tggagatcca tcagacccat 42960
atgggtcctta ttttagtaac tcttcttctt tgagtgactt ttcttcttca agctattctt 43020
tatttggaaa tattgattat ttaattaatg aaacaataaa aatttcaatt ggtggaagat 43080
gggagaattt taaatctaata tattttgatt ctatgacga atcattttca ccatcaaata 43140
aatgtcttg tggttaagtgc tcaactgtta aaactcttaa taataattct aatatttatt 43200
ttaatatgct taagggtat aaccagggtg gatttaattt aggtcttggt cttgataaaa 43260
attcatcaaa tagaaattta tattatgac cagaattttt aactaattac gaagttggaa 43320
ttaatagtaa atttttccag tcaaaaattaa atcttggagc agtcctgttt tattctgac 43380
gaaaagatca acaggtctta atttcaacc aggttgatcc ttcagatccc aacacttttt 43440
tatacttaac ccaaaatgct gcagaaggaa ttaataatgg cttagagtta aatatagatt 43500
atgcactgaa taaatctcta ggtatatattg ctaatttttg attgctaaat acagaaaaaa 43560
aaaattggat ttcaagacca gatatagaag gttagagaaca ggcgcattgct ccaaaaaata 43620
gttttttcaat aggcatcaat tggaagccaa caaccaatc ttatttatca ttgaatgttg 43680

ttggtaaaag tgagttctat tactctgatt ccataacaa tacttctgag tcatacaact 43740
taacaaacat taattatgga tatgaacatg gacaatggac ttattcatta tgggcaagaa 43800
atatttttga taagtactat tcagtaagag gtttttactt tggtaacgag gcacctgatt 43860
ttatagatac gctctacaga agacatggag atccaaggca tataaggagtc atggtccaat 43920
atgatttcta actttataag tgaattttgg atggaaatag cagctgttgt atttgctatt 43980
atttatttat tacttgctgt aaaacaagac gtaaagtgtt ggtttgcggc tataattagc 44040
tcaatatatt atttctttat tatgtatgac gcaggcttat atatggaagc ttacctgcaa 44100
attttttata tcatgatggc cttttatggg ctcaacagt ggagaagtgt cgatactgat 44160
gctcctcaat ttattgttag aacatgggat aaaaggatgc atattaagat cattacatta 44220
atagttatca tgactttaat ctctggattt ttattagaaa aatataccaa tgctatcttg 44280
cctttcatag atggactaac aacatgggga gcaatagttg cgacatatat ggttgccaaa 44340
agacttcttg aaaattggat atattgggtt gttatagatt ttatttctat ctttttattc 44400
atgtccagag gacttttatt aacctcagga ttattcttta ttatcttgt aataatatat 44460
tttggttaca tgtcatgggt aaaaaaaga gacgatatta gtgcagaatc ttcataacaa 44520

tctagatata aaaaatcata atctagagat tattaanaa attaatcag gccctgtatc 44580
tgaaaatatct atttgtaatt ttgataacat caagcaatc ttaagagttg atcatccatg 44640
cgcacacaaa ataaatgtag atcgagaaaa tgaaattttt atactcagcc aactaaaaat 44700
ttagatttt agtccagagg ttttatttag tgatttgtct tatggaattt tagtatggag 44760
atatatcgag ggtattgaat tttcacttgg taaggattcc aatgaagttt ttttaaaac 44820
acttggaacc gaattaaaa aaattcatga tattgatctc cccaaaagta agaaaaata 44880
ttttagcaat gacataaatt ttacagaaa ttactaaag gaggttcctg aaaatataat 44940
tcttcataga ggatttgatt tatacgacaa actcaataac tctgataatt atgttctctc 45000
tcataatgac ttgaataaga caaaccttct ttggagggat aggttatttt ttttagactg 45060
ggagtactcg agttttaaca atcctttttt tgataattgcc tcgttatcga atgcttataa 45120
cttatcaaaa gttgataggg caattttatg gaaagcttat acaaaataatg aatattcagt 45180
attaaatgat acaaatctta gagaatggat gcatttttgt cattatttag agtacatgtg 45240
gagtatttcg ctcatacaaa atggaaaaat tgatcagaac accctaaatt taaaaaaatt 45300
agagaaaaaa ttaaaaaata ttatttaac aacaagtatt gcgcattctt tagttgttgt 45360

tttactatta tgtaagaat ttacatttaa atgggggttt gatatgaaaa ttttatgtgt 45420
cttatatgat gatccaaaa caggtatgcc agaaaggtat gcaagagatg atttaccaaa 45480
gttagataag tatcctgatg gaatgacact tccatcccca aaatctatag attttactcc 45540
tggtgagtta cttggttgtg tatctggaga actagggctt cgaaagtttc ttgaagatgc 45600
tggccataca ctagttgtta cttctgataa ggatggagat ggatgtgagg ctgataaaga 45660
attagtagat gctgatattg ttatatcaca acctttcttc ccatattatt taacaaagaga 45720
caagatgaaa acagcgccta atttaaaat ggcaattaca gctggcattg ggtctgatca 45780
tgttgatctt caggcagcca tggataatag cgtggatgtt gttgaagtta cttactgtaa 45840
ttctcgttca gtcgccgaac acatttgtat gatgataccta tcgatgggtc gtgattatca 45900
taccacaacat agaattgtta aagagggagg atggaatata gctgatgctg tacaaaaggtc 45960
atatgatgta gagggcatgc atgttggtac ttagctgca ggtagaatg gtattgatata 46020
gctaagaaaa atgaaaccct ttgatgtgca ttgcatcac tttgatattc ataaactctc 46080
tgatgaaata gaagcagaac taaacctcac ctatcatgat tctgtagagt cattagttgc 46140
tgtgtgtgat gtagttaata ttagttgccc attgcatcct aaaactgagc acttattcga 46200

tgatgaaatg attagtataa tgaaaagagg tgcatatatc atcaatactg ctcggtggcaa 46260
gatttgtgat aaagatgcta ttgcaagagg ctagagtca ggccagctaa gtggttatgc 46320
tgggtgatgtt tggttccac aaccagctcc aatgatcac gtagggagaa caatgcctaa 46380
ccacggcatg actcctcata cttcagggtac ttcgctatct gctcaaaaca gatatgccgc 46440
tggggttaga gaaattctag aatgttattt tgcaggcgaa ccaattagag acccatattt 46500
gattgttcaa aatggtgac ttgcagggtat ggggtgcgac tcgtacaca agggtacagc 46560
cacagatggc tcagaagagg ccgctaagta taataaaataa gtttagaac ttacttagcc 46620
tttccttaat aatagagtct gcctctgcc taatgctatg cattaattct tccactgtcg 46680
ggatgtcatt aaccagtcca gcaaccattc cacatgacca ggctccaacc tccatagttc 46740
cttcatgcat aatttttggg tagactcctg caacctcatc cacaatatca gcaaaagtta 46800
attcatcacc aagagctttt tctttttcaa tcaatctctc aacagcttca ttattaagaa 46860
ccctttctgt atttgttaat gatctcatga tgagtctagt atctaactca gaagcattta 46920
cgatagcctc ttccacattt tgatgaacag gtgcattcttg agtagcaata aacctagttc 46980
ccatattcat tccctcagca cctagtgaca tggcagcaac caaactctt ccatctgcc 47040

ttccaccgga ggcaacaaat ggtatttcaa gctcatctgc tgctctagggt aagagtatga 47100
aattaggaat gtcatcttct cctgggtgtc cgccacactc aaaaccatct acagaaaaccg 47160
catcgcaacc tattgcttgt gctttttaatg agtgccttac agagggtgcat ttatgaataa 47220
cttttattcc tgcttctttt aaagctggga ggtattctgc aggatttctt ccagcagtct 47280
ctacaacagg aacccttcca tcaataatta cttttatcaa accaggatag tctgggggtg 47340
ttagtgatgg taaaaatgtt aaattaacag caaatggctt attggtcatt tctttgcatac 47400
tggcaatttc atttgctaatt ttctcaggcg taccctgtgt tagaccagta attgttccaa 47460
gtccacccgc atttgatact gccgctgcaa gctcgggcaa accaacatgg tgcattccac 47520
cttgaatgat tggatgttct ataccaaaata attcagttat tttagttttc ataatactct 47580
cctatttatac ttccattggg gttaaaaaaat cttcatattg agtttgaaagt ctttgcatgc 47640
cacttatcca tctatctctg tcattacctt ttcttttaac atatctaga acctcactat 47700
gaggggttac taaaaaacctc tcctcctcta tggcggttaag aacatccttt gcaacaatat 47760
cagcttccat cattccatca actcctgcta caccaggacc atttgcatc atagcagttc 47820
taactgcttg agggcataaa caggaaacgc cgataacctt atttccatat gtaattttta 47880

tccattcagc aaaactcaca gctgcagctt ttgtaactgc atatccagcc gcacctagtt 47940
gagttaaaag accggctgct gaggatgtat tcataagata gccttcacct tgttctatca 48000
tttgaggaag cacatgtttt gcagcatgaa tatgagactg aacattttaca cccatatca 48060
tgtcccaatc cgaagtatct gcttcaaaaa atcctgggctt tccaccgata cctgcatttg 48120
aacaataat atcaatacca ccagaaaatt cattagcctt ttgtataaca ttataatgt 48180
cgtttttcttt ggaaacatca gcacttactg caagcccatt aacactcttg gctgtttcct 48240
ccgccccatt taaattcata tcaacacata ctatagaact tgcccagat gcatagaact 48300
cttcacataa agccttacca attccactgg cagcccctgt aaccacaact cttttattat 48360
taattttcat aaattgacca cttttttaat tttttattta actttattgc ttagatagtt 48420
tattacaaaa caacaatact taatattgca aaatcgcata cccatcttta taaattttatt 48480
gctattatta gtgctagaaa atgaaatact gagatttaat atgaagaatg ttgttggttat 48540
tggctcgtcc ggggcaatag gaaaagcctt tattgatagc tatatcaaag atgatgatgt 48600
tgaaaatata ttttcatttt caagaacagg cttttccatt gaggataaaa aactccatag 48660
tttttttatt gatattgagg atgaaactag tatttgtgat gccgcagaga agatagacaa 48720

gtcctcaata gatgaaatta tcgtcgcaag tggaatactt cataataaag attttgggcc 48780
agaaaaaagt attagagatt taaatgcaga taacctttta aaggtcatta aggttaatac 48840
tatcggccca acaattgttg gaaagtattt cattccattg ctaaataaaa agaaaaaaag 48900
cgtcttagca tttttaagtg caagagtcgg cagcatttct gataataaaa caggtggttg 48960
gtatgcctat agagcgagta aaactgcact taatcaaatc attaaaagtt ttagtattga 49020
attacgaaga accaatccaa atgccattat ttttgggtcta cagccaggaa cagtagatag 49080
tgaattaagc gaacctttta aaagaaatgt aaaagaaggt aatttattta ctccagaata 49140
tagtgtattg cagctaaaaa atattattga tacagcaagt ccatctgatt caggcaaaact 49200
aatctcttgg gatggggaag agattcagcc atagttgggt atgaatatat tttcatatta 49260
gaataaatte ttaggggaaa aaatgagtat taaatattat gactgggtcta aatttcaggc 49320
caacactagg ccaaataaag ttgccataag agagctagat aataacaaga tctatactta 49380
cggagaattg gataaaagat catcaaggct tgcatcacat ctccaaaagt caggaataaa 49440
aaaaggagat cgtattgcga tactatcgct aaattgttca gaattttttg agctagaatt 49500
tgcttgcgga aagattgggg caatagagat accattaaat tggagattaa caaaacccga 49560

gctcagttat attcttaatg atagtgagcc aaaaactcta atttatgaca atcagtttga 49620
agaaatgggtg aaagagctaa aagaagaatg taatatctct gaaatcatag ctcttgatca 49680
atttgaccaa gaaagtgatt atgaaaaagt tttagtaat gcttcaggca tttattatca 49740
ggaagaagtt gatctagaag ataacattat gattatgtat acctctggaa caaccgggtca 49800
ccctaagggc gccatgatca cgcacaaaat gcagcttttt aatgttatta atttaggtat 49860
ttcagcagct gtttcccctg aatcagtcga tttagttgtc ctccctttat ttcatacagg 49920
cggaatgaat tgttattcaa atccaattct tcatgcaggt ggcgagtaa tattacttaa 49980
agagtttgag cctgggaaag ttctatcaat tatcggcagc tctgactatg gagttactca 50040
tctgtttgca gttccagccc cttatcaatt tatgatgaat catccagatt ttgaatcaac 50100
aaatttatca ggagttaagt atgctggagt tgggggcgca ccttgtgcag aggctatttt 50160
gaagacttat ataagcaagg gtgtttcgat gcagcaagga tgggggtatga cagaaaactag 50220
tccagggtgct actgggtcttg aatcgtccga ggctgaaaga aaaataggat ctgctggaaa 50280
accgcttctt catactgagg tcaagggtgt tggagatgat gggaatgaac tgcctgctgg 50340
agaagtaggc gagatttata ttaaaggccc aaatattaca cctggctatt ggaagaaaga 50400

agaggctact agagattctt ttgaagacgg gtggttaaaa acagtgatg ctgcttactt 50460
tgatgacgag ggttttttat acatagtga tcgatggaag gatatgtata tctcaggtgg 50520
agaaaaatgtt tatccagctg aagtttsaaa tgatatctat cagttaccac aaatcgcaga 50580
agtggagtt attggtattg atagccctaa gtgggtgag actggtaaag cctttgttgc 50640
tttaaaagccc gatcatgaat tgacggcaga agaagtcata gatcattgtt taaaaaatct 50700
agcaaagtac aaaattccag agaaagttga gtttattgca gctcttccaa gaaatgctac 50760
aggtaaagtt ttaaaaagaa cattaagaga tatgtaatat ttttaagcca aaaaaaaccc 50820
agcctaagct gggtttttta tttaactatt taaagttatc taatgttaag cattagaaga 50880
ttctttaaca gcaacattcc atataattaa accaaataga atcttgtaa caaagtcagc 50940
aaggttatag ataaggttta agttaagagc tgatccaccg tcacccatca ggtaacctgt 51000
gaaataaacct acaggataaa tcgcccacc aaagatgata atatacatca ttgtgttgta 51060
agctgattgc acagcaggac ttgcagtatt acatgcagat ttctcttctc cagcccataa 51120
ttcataaatc atgtataccc aagctaaaca cccaataatg aatgcaggcc atgcagccat 51180
gattcctgct tcacccatgt aaccaaacac aagcataaca agagaaccaa ctagtaattt 51240

cttaaataat gatccagcaa cattagttgc agcagcaaga attaagtaga attcacatat 51300
taatagagga actgttagta accaatcaat gtatctaaat acagttggcg aatcaccagtg 51360
ttcaatccat acccctctca tgtacatgta atgccagaaa gcaataccag taacaagacc 51420
agatacagtt aatgatgttt tccattttgc agaaactcta tctctttcaa caaagaaaaa 51480
tacagtagat gctaataaag cagcagtaac taacccaaaa gaaacaccag tgtaactact 51540
agcatcaagg tcaccaccac ctgcagcaaa tgtaggaagt gcaataaac tacctaatat 51600
cagtaataat ttcatatata actcctatat taatgacaaa gaagcataat tgcctccccc 51660
ctaattaaat atgaaattta atagtaacca atattaatag aggttaaaaa taaaaagcat 51720
ctattttttt aatagaaaaat atgtgtaata aaacaccata ctttaaatata atattttatt 51780
tgcataagta gacttgttta gtatcatatt taacatgaaa gttgcaattt atcctgggttc 51840
ctttgacccc atcacaaatg gtcacacgga tattattgat aggggttgcg gactctttga 51900
caaggttgtt gttgcaatag ctaagagtga atcgaaaaac cctcttttta gcctagagga 51960
tagaattaat ttagcccaat ctatttttta aggaatatgaa aaagtagagg ttgttggttt 52020
tccaagaaaag ttaacagttg atcttgcaa agactatgga gcttggtgcaa ttataagagg 52080

cctacgagca gtttctgatt ttgaatatga atttcagtta gcaacaatga ataggctcgct 52140
ggctcccaat attgaaagca tttttttaac accaaaaagaa agtctcattt atgtatcttc 52200
tagcttaatt aaagaaatat cagacttaaa aggcgataata tcgaagtttg ttcatccttat 52260
agtagagcag gcacttcgag cgagtagaca cttagctctg acaagccttca caaaaaaaag 52320
tagctctttg attgacgac gtttttataaa tagtgccttt gcatttattg cagggctcgc 52380
cttctcttcc atatacattt aatttaaatt taaaatatcc tggactacca tcagctgagt 52440
agaagtcttt taatgtagtt cctccaacct ctattgcttg ctctagtatc ttttttccag 52500
cagctactaa tcttttataa gcactctaaat cttaactcatt ggcatttttt agcggatgaa 52560
ttttagctaa gaaaaggctt tcagacgcat aaatattgcc gataccgact acatttttttt 52620
gattcattaa atagctttta atatttgttt tagagtgact gcaccctgaa aaaaaatcct 52680
ttgcattaaa attttttagaa agagggttcag gcccaagggt ttttattaac ttatgtttat 52740
cgatatcaga agtaagatgc attgatccaa accttctggg atcattataa ataattcttt 52800
cctcatcaaa aattaattca atatgatcat gtttgataaa gaagttttca ttattttttg 52860
caattctaag actcccagac attcctaaat gaagaataat ttttttatca ttagatagtt 52920

taaaaattat atatttttgct cttctttcta agcttttcgac aacttgatct tttactgagg 52980
tttcaaagct atccaccact ttccatctaa ggttttctgtt gtgtatcctt gcttctttca 53040
gcaaagagcc tttaaatttt ttaattgctc tcagggttgt ttcaacctct ggaagttcag 53100
gcatttttaa ctgagtaaat tattaatttt tactatgtca gctggagtaa ttgtttcccg 53160
tgcgagccct agtcttagat tggcaagtat gtaatcatat tttgcattag caagattttt 53220
ttccgcgctg tataagtttt tttctgcctg caagagatca acaacgtttc ttgttccaac 53280
tctgtagcca acttgagtcg cttccagggc actagtggcc gaaatcactg cttgtttttg 53340
agcattttaca ttgcaacta atgttaaaac atttgaaaac tgggatctga cttcttgaat 53400
aatcccttctt tctgtaaata gagtattttc atttgctctt tcatactgtg aatatgcttg 53460
cttcccttctt gagttaacgg cgccgccttg aaagagtggc atacttagct gaattgcata 53520
attcccttctt cctgttactg atggaactgg aataccttgg ccattgatat taaaaccttc 53580
atagttaaat tggtttgttt cagatttctga ctgactttcca acaatgtcta tcttaggtaa 53640
atgatttgaa gctacacttc ttgcactgct tttcgctgct tttctttctca aatatgctgc 53700
ttttaactgg tagttatttt ccaatgctaa ttcaacccat gtctcttttg aacttggtgt 53760

tggcagggtca ataagcaaac catctcccaa ttcatTTtaag ctgaatatTTt ctctaccaat 53820
cagagcattt aaagactctc ttgcagaata aagtgatect tctgttctaa ttcttgaggc 53880
tttacttaga tcaaatgcc aattgagcctc ttgaactcca gttatggctg ataaccacaac 53940
atcgaatcctt tgttttgctt gatcaagttg ttttttaata gcttttttctt cagatatTgc 54000
tgcattttaga ttatcaatag ctctaagtac gccaaaaataa agctcagcag ttcttactaa 54060
aagattttgc tgctcaaatg caaagtctgc ttcagcagca tctgtaagag atttagattg 54120
cctatatTga aaccatgtat ctagtctaaa gagcggctga gtaaccctgg cagatgtaga 54180
aaaagagtta tattgctgct gcagtttctt gttttgatag tattcgttcc agttagttga 54240
tccactcaag gtaatactcg ggagaagcg agctcttctt tgaaccctaa gctctttatc 54300
tgctaaatat gaatatccg ctgctttata tgtagggTca ttctcaagcg cttcattata 54360
gatatccaaa agactttcag atgaaatat aaatgagata aaaagtgcta aaaatgattt 54420
tgtgtaaatt ttcataactt atttttaacct attaataatg tttgcagTgc aaacattttat 54480
ttttttttaat tttattgaat ttattttata tttcttagag tagaatatct ctaacaagtt 54540
caataatttt ttataataaa aaggctttta aaattggcta aaaattcata tgacgctcag 54600

gcaattgaag tcttatctgg actagatcct gtcaaaaaaa gacctgggat gtatacggat 54660
acatctaacc caaatcattt aattcaagaa gttcttgata attcggttga tgaagctctt 54720
tcaggttatt gttcaaacat aaaagtatct gttcttaaaag atggcctttat taaggctctt 54780
gatgatggaa ggggtatgcc aattgatgag caccgggaac ataaagtttc aggtgttgag 54840
ctcatccctt gtaaaccttca tgctggagcg aaattctctg gagatgatta taatttctt 54900
ggtggccttc atggtgttgg agtttctgtt gtaaatgcc tatctgatga attagaggtt 54960
agagtaaaa gagattctaa agaataccaa atcactttta ataatggaga taagtcttct 55020
gaattaagc caattgggga agtggggctc agaaattctg ggacatcaat taaattcaaa 55080
ccaaatccta catatttga aactatagag attcagataa aacagcttaa gcatttatta 55140
aaggccaaag cagttctctg tcctggatta acgatagagt tcgttaatga aaaaaagact 55200
gatgataaac aaaagtggta ttttgaagat gggctcaaaa gctatttgat tgattcttcc 55260
gagggagcag acttgggtttt gctagattca attgtatgct ctaaaaaatc tcaagctcaa 55320
gagcttgaat ttgcaatcaa ttggtcata agaccccaa aaaataaact cgatgaaacc 55380
tatgtgaatc tcataccaac tgctcagggt ggctcacatt taaatggctt taaggctggg 55440

cttttagatt cattaaaaga attttgtgaa tacagaaatc tattgcctaa aggttttaaa 55500
attaatgcag atgatgttct taataatgca atttttataa ttcatctaa gcttcagaat 55560
cctcaatttg cagggcaaac caaggaaaga ctagattcaa aagatcacat gtcgttcgtc 55620
tcaagtacca caaaagacat tttaagtatt tggcttaaca ctatacaga agagggcgaa 55680
agaatagcag aacttgcaat tatgtctgct cagacgagag caaaagtctc caatatagtt 55740
gaaagaaaaga aaacttttag agggccagcc ttacctggaa aactttcggg ctgtaatagt 55800
caggacttaa atgaaacaga gcttttttta gttgagggg actcagccgg aggtccgca 55860
aaacaagcaa gagaaagatc ttccaggca atcatgcctt tgagagggaa gattttaaat 55920
acttgggact tagaaagtgc agaaataata aaatctcagg agataaaaaa cctatcaact 55980
gcaattgggg ttctgccagg aaataatgac ctttcatcac taagatacgg aaaaatttgt 56040
attcttgag atgctgattc agatggctg catattgcaa ctttactttg tgcattgttt 56100
ctaaggcatt ataaatcttt agttcaagag ggaaggatat atatttcaat gcctcctcta 56160
tatagaatcg attctggtaa agatgttcta tatgcacttg atgataaaca gcgagatgaa 56220
atagttactg aatttaaaaa gaagaagggc aagcctaaag taaacattca aaggtttaa 56280

ggacttggtg aaatgaatcc acctcaacta agagagactg tgatggaccc tgctactcgt 56340
cagcttggtc agctttctat cagctcaagc gataatgcaa attctatgat ggacttactt 56400
ttgtccaaaa agaacgcacc agcaagaaaa gaatggcttg aaaagaaaagg gtctctagca 56460
aaaataaaa tatgaaagaa caaataacct caattagcct caagcaatat gctgaagagt 56520
cttatcttaa ctatgcaatg tatgtcattt tagatagagc tttgcctaata attggagatg 56580
gccttaagcc tgttcaaaga agaatactct atgcaatgtc agagcttggg cttgatgctg 56640
gctcaaagta caaaaaatca gcaagaactg ttggagatgt tataggaaaa tttcatcccc 56700
atggagacgg cgctgcatat gaagctatgg ttttaatggc tcaaaaatttc tcattcaaat 56760
acccttttgt agatggtcaa ggtaactggg gttctcagga tgatccaaaa tcttttgctg 56820
caatgaggta tacagaatct aagttaacta aatttgcaaa tcttttaatc tctgaattga 56880
agtctggaac agtcgattgg cagcctaatt ttgatggctc tcttttagag ccagtaattt 56940
ttccagccaa actcccatct atttttattga atggcacttc tggaattgct gtaggaatgg 57000
caacagatat tccatctcat aatattaatg aaattattga tgccacagta catcttattg 57060
ataatccaaa atcacagttg gttgattttac tcaagataat taatggctct gatttctcaa 57120

ataattcgcc ataattgct agcaaagatg agctgaatga aatttatctg actggaaaag 57180
gcggcttcaa agctcaagcc caatgggcgc aggataagaa tcaaatcatt atcaacgcat 57240
taccttatca agcatctggg tctaaaattt tagagcaaat agctgatcaa atgccttaaga 57300
aaaaaattcc aatggtggtt gatcttactg atgaaggaga ccacaaggag ccagtaaggc 57360
ttgtcataac tttaaaatcg aacagagtaa atgctgaaga tgtaatgaat caccttttttg 57420
catcaactga ttacaaaaa aattatagag taaatatgaa ttgtatttca ttgaaagggtg 57480
gaccaaaagt tttctcccta gttgatattat tgaaagaatg gctagtcttt agaaaagaaa 57540
ctgtaataag aaaactagaa catagactcg accaggtaaa cgataggctg catatccttg 57600
aggggttatt aattgtttat ttagatttag ataaagtaat aaagattatt agagaatcag 57660
atgagccaaa aaaagatat attacagctt ttaaaactttc tgatatccag gcaaatgcaa 57720
tccttgagat taggttaagg caactagcca agctagaaca aatcaagtta gaacaagaaa 57780
gagatacttt ggtagcagaa caaacagata ttgaaaaaat actaagctct aagacaagac 57840
ttaaaacatt aattaaaaat gagcttattg aaataaaaga tgagtttgga gaagtaagam 57900
aatctccaat aagagaagca acagaagcca aagttttttc tgaaagaaga actcttgtca 57960

ctgagcctat tacagtagtc ttgtctgcag ctggctggat cagaagtgca aaaggccatg 58020
agatagacc cagctcactt tcctatagag gagaagatgt acttcaagat tatggaagag 58080
gaaagagcaa tcaagtttca gtttttcttg actcaaatgg gaaggcttac tcacttgcaa 58140
gtcactctct tccatctgct agaggaatgg gtgatccctat tacaggaagg gtatctgcag 58200
attctggagt aaagttcatt tcttcattga ttgggaatga tgaagataaa ttcatgatta 58260
tgaatactgc tggatatggc tataatttcag agtttaaaaa tatggtttct aataagaaat 58320
caggaaaagc atttatgaaa atccctcatg aagcagacct tcttaaagcc attaaagtaa 58380
gagacgatca tttgtatata gcagcagttt caaatattgg cagactttta atttttaaga 58440
ttgatgaatt gccaaactctt ggaaagggca aagggaataa aataataaat ataccaaccg 58500
ctaaatttat agcaaaaagaa gagttaatga cccatgcgca acttggttct gaggctagct 58560
ctttaaggat tgaaaagcga aagagattcc tcactttaaa actcaaagat ctagaaaact 58620
atatattctac aagagcaaaa aggggaaata tgcttcaca aggatatagg aaggtagata 58680
aaatgattga agaggttgag ttagaagtta aagaagactg attatagatt tctcaaaaat 58740
ttttaatccc tcataaacta gctcatttc tattattagg ctggagaaa acctaacagt 58800

agatgcattc gccttaagaa tcattaaccc attatgatgg gattttttta ttaaatcatc 58860
tatttgaatt ttgctatcct tgcttacttc cactccgacc cataaaccag cagaagtatat 58920
ttttttcaaaa cattttgtgct tctcatattat tttattcaac aaattgataa atcctaacttc 58980
tttcttttaaa acctttattta aaaatgattt ttttagaaatg gtatctataa cttcattacc 59040
tatggcacat gcaattgggc ctctccaaa agtcgtgcc aagaccctg ccgacatatg 59100
ttttgaaact ttatctgaag taagaatgcc tcctagtggg aagccattag atataccttt 59160
agcaaaaaca agaatatcag gagtaatgtt aaattgctca taagcaaaga gtgttccagt 59220
cctgccaatc ccggattgaa cttcatcaat tatgacaagt gccttatgtt tttttgcaag 59280
tttcttttatt ttgctataa atttttttatc tgctttttgtt attccagatt gccattgaac 59340
cagtttctaataaaccgctg cagtctttatc tgaaaaaac ttttctagat tagtaatgtc 59400
attatatgga tgattttttta tccctctagg cagaggcgca aaaccgtctg taagatgctt 59460
agcttttagct aaagctatac caagcatagt tcttccatga aatgacgttg aaaatgaaat 59520
gacctcattt ttatttttat taaactgtga tgagcaaaaat tttcttgcta tcttaactgc 59580
agcctcaatt gattctgccc ctgaattgca aaaaaatact ttatctgcaa aagaattctt 59640

gcataacttt cttgccaaag taactgaggg ctcatattatg tagagatttg ataaatgcc 59700
aagctcctca gattgctttt taagaatttt tattagatct ttattgaat gaccaagggtt 59760
agtgacggca atgcctgctg tgaaatcaat atatttctta ttatttaggt ccatacatg 59820
cgatccactc gcttttttta caacaaaatc tgcaggagcg taaaaaggca ccataataat 59880
agtatatctt ttaattatct ctttcataaa tatattgtat cgtggtttca gttcaaaaaat 59940
taaatgacca aaaatttcta gtatctttaa aaccaaacag ctactcatt ggctttaata 60000
gaataatatt tatacttagt atatctattg tatcgggtgg catagcattg atatttttct 60060
tttttggagc aacgctcatt ttgccttttg ctggtttaga gcttgggtatt ctattcactg 60120
cattctattt aagttttaaa tggagtgata aaaaagaaaa aatatttatt tctcaagatc 60180
ttgtaactat agaaaaaggc tctaattatg ctgaatatata atgggaagag ttcagggtcat 60240
ttacctcttt ccaggtttca aaagacagaa gagatcttct taagctaagc tttaggttcta 60300
aaggcgagga tgttgagggtt ggcagctttc taaatgaaga tgataaaaaat gtattaatag 60360
aagagctaac tcagatcata gatacatata atcacgattc cttctcaag ccagagcgttt 60420
aatttctttt taacttcttt tagctttaat tcttttattt cagggatgcc atttctaaaa 60480

tttggaatg attctccctg tattaatggc ataagatatg caattgcttt tggggtaacg 60540
tccattccat ttttagctat aaaagatttt ggtaattttt tttctaaatt agctattttt 60600
gccagtggag ctggttcaat ttccaccta tatttttttag ctttacctct aactattatg 60660
ggcataaacac cattcatgcc ttctactgca tactgaactg ccttggtccc aacagccata 60720
gcttgctcta ggtctgtttt ggaggcaata tgctgtgcac ttctttgtag atagtcagaa 60780
acagcccaat gatttttttag ttttaattta tcagtaatta agttggcaat gtaaggcgca 60840
actccacca gttgagcatg accaaatgca tctattgttg ctgattcaga aagaaatctg 60900
ttcttattat ttttcaacc ttcgatacc acaacaacac agtagccatt tttttaaca 60960
acacttttaa cttctgctag aaattttttt tgattgaagg ttatttcagg taataaaatt 61020
atatgaggtg catctccttt ttcttttctt gcaagagaaag atgcagccgc catccagcca 61080
gcatgcctgc ccataacttc taatataaaa acttttgttg atgttgctgc cattgaccga 61140
acatctaatag aggcttcttag agttgatgtt gctatatatt tagccgccga cccaaatcca 61200
ggacaacagt ctgttaatac caaatcgtta tcaacagttt ttgggatggc tatacagtta 61260
atgggataat ttaatttttt acttatctgc gaaactttaa atgcagtatc agctgaatca 61320

ttaccgccat tgtaaaaaa atatcctata ttgtgcgctt taaaaacatc aataagcctt 61380
ttgtactctc tttcgcttga ctcaagatct ttttaactta acctacatga cccaaaagcc 61440
ccaccaggcc tataatttcaa tgactctaat gcagatatg attcttttga tgtatctatt 61500
agttcctcat tcagcgcgcc caatatcca ttcttgcctg cataaaatctt accgatgtct 61560
ttatgcttct tggccttctaa aatcagagcg ccagctgtag catttataac tgcggtaacg 61620
ccgccggatt gagcataaaa cgcattttttt ttcatttttt ctccatcaaa tgtactgaaa 61680
taaaaatatg taaagtatta taactaatat gcgaatacat atcttagga tctgtggrac 61740
ctttatgggc ggccttgcta agatacttaa agagtcaggg catgaaatat ctggatcaga 61800
tattcaattt tatcctccca tgtcagacta tctgatagc tttgatattg aatgatcaa 61860
gggctatgac ataaaaagca tgccagatgc tgacttgat gtgattggta atgcactttc 61920
tagaggaaat gaatccgttg agcatatttt gtccaattct ctccggttta agtcaggccc 61980
tgaaatgctt ggagaaattt taaaaataa aaagggttta gctatatctg gaacacatgg 62040
taaaacaacg acttcataata tgttgactca tattatgtta gatcagggtta gggatgtcgg 62100
ttttctagtt ggtggaatat caaataatat atctggctct gcatgtcttg ggtcagatgg 62160

aactttttgtg attgaggcag atgaatatga ttctgcattt tttgataaaa gatcaaaaatt 62220
tattcactat tcaccaagca ctatagtcac caataatatac gaatttgatc atgctgatat 62280
ttttaataat ttagatgaca ttaaaagaca attccatcat ctaattaaaa taattccatc 62340
aaatggaaat gttgtttatt ttgctgatga taaaaatata agagatctta tcgatatggg 62400
aatttgggtca aatcagatag caatcaataa taatgctcat tcaattgagg cagttttattc 62460
tgataagact ctaaaatatg aagaaagtat ttattcatta aatgagttac ctttaatatag 62520
agaacataat tttaaaaact acatttcggc tattttggcg gcaaaagacag atggaattca 62580
aattcaggat tctatcaaat cattagctag ctttgatggc gtaaaaagaa gattagaata 62640
taaagggaagt tttgatggca tagaaattta tgatgathtt gctcatcatc caaccgcaat 62700
agaattttct tctaattgcc tcgtaacaca aaatccatca aaaaaaatc ttgggtctcat 62760
tgagctaggc tcaaacacta tgtctggggg ttctcatggc ttgtcttttg ttgaatctgc 62820
aaaatcttta actcatgtta tctggctaga tcgcaataat gttttgtcag agaattgctag 62880
cattgaatct actaacacta ctgaagatth tattttctgca gcgatatctg cttttctcaga 62940
ttatgatatt gttattttga tgaccaataa agacagccaa aaaatatata aaccattgt 63000

agatcacttt gaaaaataat aattaccag ttttccctt aggaatagtc gccctcccag 63060
gtagcatcca atctcttcaa attttgagc ctaggtatat acagatggtt aaacatgtc 63120
tatccaagaa ccatggattt gtaattgttt ttaatgccaa taatgagtct caaggcgatt 63180
tcactttttc taagaaggga agttttgttg aaattataga ctttaataat ttgccaaaatg 63240
gccttcttgg gataactgta aaatctataa ataagtgat aatcagtaat atatgtcaat 63300
tagaagatgg actgcataat gctgatatta aggcacagat agatccagag gtagatgac 63360
aagctgtttt ggcagaatat cctgagatat ctgacattct ttctcagctt gtaaagcatc 63420
ctaagattag tgacctgccc atccaggttg actttggctc tgctgattca gtgcatacc 63480
acttagcagg ccttatacct ctaagctcaa atgagaaaca aaaactatta gaagcattcg 63540
atgcagcaca gcggatgaga attctttcag actatatatga aagaatatct actacataaa 63600
ttatttattt taatattatt ggcggccttat tagatttttg tctaagccaa ttgatggact 63660
taaaaaccgt agggatttct acaaaccctt ttccaatata gtatttgcct ggatagttac 63720
taagcattag cccaataata atagtgaaca aacctggcca ggtaacacca gcatcatgat 63780
tcccccaata ataagggaat aacctaaaat attttaatc agtaaaaacta aataccacaa 63840

caacgggtta ttggtcttaa atttcgagac atcttttcttt ataaaaatag cacttggaat 63900
taatccagca agccacctca tactaacgag actaaaataa aaaataaata aagatatatga 63960
gctcagccaa agtataagaa ctttataaga ttcaaaaaaa attatgatgg catttagggt 64020
attaaattcc ataattaccc acctttttat ttaataatga tttagtatag caatgtcttaa 64080
aaataaaattt aaacattaaa taaatatata ataattttgc caaaaaaagc cttaaatgat 64140
atttgcactg aaaccctcag tgataataaa gccgaaaatg ttctatcact tgatatcaaa 64200
ggatatttctt ctttttgctga caacatcatt attgcaactg caaatcgaa taggcatgca 64260
aagtctttat ccgaaaagt agttgaggag ataaaagcta ataaaattag tatcatgggt 64320
gtagaggga agacagaatc aggttgata ttagtagact gtggtgaggt tgtagtaaat 64380
attatgaaga atgacataag agagttttat gatttagaag gtctttgggg tgaaaaacacc 64440
ctcatcgatt cttcgaaata aatgctatta aatatcataa gtgtttggaaa ctcaccttca 64500
tcttgggagt taacaggaat agaattattac acaaagcaaa tccctaaaga agtaagtcta 64560
aattttatta acgtaaaagg gcagcaaacac ccaaaaagat ctacagaaga ggtcttaaaa 64620
cttgaatcta agttaatctc gtccaaaata gattctaagt gatataattgt ttgctgggac 64680

tcaagtggcg agtcgttgaa taattttgaa ttttagtaaat tttttgaaaa atccatgctg 64740
gaaaatatga agctttactt tgtaattggt gggttcatttg gcataacctcg agatattctt 64800
gataaatcta ataagataat ctcgatctca agtctcacc ctaacctcacag gctttttcaag 64860
attgtttctta tagagcaaat ctatagggtca ttttcaattc tcaaaaatct tccctaccac 64920
aatgattga tctgaatgag aggtttgttg aaaaaagaag ctttttcaat agattgttgc 64980
ttatatattt tttctttggg gccctgtttt tattcttgct tttcaagacc tattctttac 65040
aggtttctag ctattctgac tatgaactag cagcactaaa gaataaaaca aaagaagtgc 65100
tagttcagcc tgttcgagga gtcatctatg acaggaatgg aaatatctta gttaataatg 65160
tccctacata tgacctaaata atccaaccat caaaaattaa aaacttagat gattttcata 65220
ttaatatctc taaaataata gagctatcag attcagaaat tgaaaacatt attaaaaatt 65280
ttaaaaggag tgcaacttat aatagagagc taacaattaa aaaagatctc tcgaaagaag 65340
aaatagcaaa atttgaagta agaagtacc aatttcagaa tgcatttata gatgttaggt 65400
atagtagaga aaataaatat ccttacttat tttcacacgc actcggttat gtaggtggtg 65460
taagtaatga taaggtttta tcaattttta aaaatcaaga tttgaagcaa tctgaaacaa 65520

cttttaagta ttcagggtggt tttattgctg gaaaaacagg attggaaaat atttatgatt 65580
cagctttaag agggtccttt ggaaaaaagt tatttgaagt tgatgcaagg ggaaggcctt 65640
taaaagagct aagttttgaa aaacctatta atggaaaaag ccttttact catttagatt 65700
taaattcaca aaaaaaagcc ttgaacaaa tgaataatag aaggggtgct gttgttgcct 65760
tagagcttaa atctggttct atagttacct acctaaagcac tccaagcttt aatgtaaatg 65820
gtcttttctaa tggcatgtcg tcggtagagt tcgaaaaact gattaatgat gtagacaagc 65880
catttttttga tagggccggt caaggtcggt actctccagc ttctacaata aagccagcaa 65940
tagcattgtt tggattataa gaaaaaata tagactggaa cttacactt aaagatcctg 66000
gatttttttgt attaccagag gatcagagga ttatagggg atggaaaaa ggaggtcatg 66060
gaacaattga tatgaagggt gcaatcatag aaagtctctaa tacttttttc ttttctcttg 66120
cttataaggc tgatattaat aatttaata gccatctttc tgagtttggg tttggcagga 66180
atgttttgtaa agattgtttt aatccagatt ctggattatt gcctacgcct gaatggaaaa 66240
tgaataaatct taattttggc tggttttaaag gagataccgt taatttaggt gttggtcagg 66300
gctatatgag tgccactcca gttcaattag catattactc tgcatttctg gcaaaaaaag 66360

gaaatcttca agagctatca ttgtcgaga gtgacagcct aagcaataact gcttttataa 66420
ataatttaag catagataac tcagactggg atcaaattca ctcaagcatg attggagttta 66480
ttgaagatcc gagaggcact gctaaaagat taaaaccttt aaaatcatat gttgttgctg 66540
caaaatcttg aacgggttgag cttgtcagca cacaacaaa ggaagattac aaactagtaa 66600
ggcaaaaatat tggtaacaga gaccatgcaa tcattgtggc ttttggcca atgcctgac 66660
cagagtatgc agtaagtgtt gttattgaaa atggtgaaag cggcggttct gttgcaggtc 66720
ctgttgctat tgctgtttta aatgagctta tcaataaatg aagaaaaaat tagacttttaa 66780
aaactttagc atttattttg atcaatatat atttattgcg ataaccctgt tgcgggtcat 66840
gggttttatt tttttatata gcgcatacaca agaagatatc agcactgttg ctaagcaagc 66900
tgtattcgtt ggtttttggt tgctgtttaat gttttagtt agccaacctg accctgattt 66960
ttataataca ttttctgggt tattttttgg ggggggastt gtattgattt ttctaactat 67020
gatttttggt aaagaaaata atggagccaa aagatggctt gattaggat tttttaccct 67080
gcaatccctc gagattatta aaatatcatt gccaattttt ttatcatcat atttatataa 67140
taagccactt ccaataagca ctaggcatc tttttattaca ttgatattaa taggttttat 67200

atttgactt gtagctagac aacctgactt gggaaccagt ttagtagtat ttatgtcagg 67260
gggttatgta ttatttttag caggattaag ccgccgttta ttggatctgc aatagcctta 67320
tttttattat ccttacctbt tttatggaat aacttttctag agccatttca acaacagaga 67380
gtttttaactt atttgaaccc agacgcagac ctttatggta ctgcatggaa tataactcaa 67440
tcaaaaatag caattggatc aggcggaatt aatggaaaag gttatcagga gggctcccaa 67500
gcccatcttg atdttttacc agagacagaa acagatttta ttttgctgt tattgctgaa 67560
gagtttggtt ttattggagt ctgtattttg ttatcagtat ttctctttat atkactcaga 67620
tgtttatatt tagcatttaa tgcaagagat agatttttgca ggttaactat aggaggccta 67680
agtttagttt ttgcctctac attattttatt aatttagcaa tggttgttg tgtagttcct 67740
gtagttggta tgcctcttcc atttatcagc aaaggcggct catctttgct atcctkttat 67800
atagcttttg ggattataat ttctatggca acacataaaa aattaatgca aagatgaaaa 67860
aaattatatt tataacttta atattcacta ttccaattac tgcggattat tcgaatcatg 67920
aagatagcca aatggtaata aatgaactcg tcacaaaaaca cggttttgag gaatcttatg 67980
ttactgcaat cttaaaaaat gcaaaaaagc gtgatgagat gcttaaatct gttgctaatac 68040

ctgctgaaa aacaaaaaca tgggatgaat atagggtctat ctttataaaa acaaaaagag 68100
tttcagaagg caaaaaattt ataaaaaaaa atattaatgc tttagagagg gctgagaaaag 68160
aatttggagt tcctaaagaa ataattactg ctatttttagg cgtagaaact aattatggca 68220
gcaataaagg tggatacaga gttttagata gtcttactac tttaggtttt gatgaccgcg 68280
gaagatctaa tttttttaga agagaaactta tagagtctct ccttttaaca agagaaaaata 68340
atttagatat caaaacaaca aaaggctcat atgcgggagc catgggatat gcacaattta 68400
tttcgtcaag ctaccgagcc tatgctgtag attttgatga agatgggttac gttgattttat 68460
ttaattctgt tgatgatgca attgggagta ttgcaaatta tctttatgtt catggatgga 68520
agagagaagg aaagatcgta acaaaaactt atccaaaca tgtagaaaa tttataaac 68580
ctcatgagtc tctaacaagg ttcatacctt taatctttaa tgaagatgga aaagatcttt 68640
tttttattgg tgatgataac tttagggtta ttgctaagta taatattagc gatgtctatg 68700
caatggcggt ttattactta tcagaagagt taaaaaaatg aaaaaattat tattcacatt 68760
actatcaact tcaatattta ttcaagcaca gagtttttgtt ccagattctc ctgagtttaga 68820
tctgaagagc tatatcctaa ttgagccaaa caccaatact gttattgcag aatttaattc 68880

agatttggaa atagaaccag ctagcatgac taaaattatg actagctatg ttgtggctga 68940
tcagattgca aatggtttaa tatctcttga tgacgaagta ctaattagtg aaaaagcatg 69000
gagaatgcaa gggctctaaaga cctttataga agctggaaaa aaagttagcg tatctgatct 69060
tttaaaaggt attatgattc agtcaggcaa tgatgcctct gttgcaattg cagaatatgc 69120
aggcggcact gaaagagggt tcgttgactt aatgaaactct tatgctgcct ctttggagat 69180
gaataatacg atcttttcaga attcaacagg gcttccagat gaaaatcatt tttcatcagc 69240
aaaagattta gccaatctga cagctaacta cattaataaa tttccagaag aatatgcttt 69300
atataagcag aagcaattta cttttaacaa tattaagcaa ttaaaatagaa ataaagctttt 69360
atggagagac gactcagcag atgggtgttaa aactgggtcat acagaagctg ctggatactg 69420
tttagttggc tcagcaaaaaa ggggcgggtat gaggccttatt acagttgttg caggaagtaa 69480
gtcagataat gatcggttttt tatcttccca aagattactt gagtatggat ttagatttttt 69540
tacaacacag aagatgttaa gtgcagaaaa agaataatcag aatattacag tttgggggtg 69600
acaagaaaaag atacttggcg tgggagttct agaggatata tctattactc ttcctagaac 69660
aagtttcaaa aatgttgaaa ctgtttataa agttaataac aatatccagg ctcctatcgt 69720

agttggtcaa aaagtcggga ctctagaaat aattagtaat gatgagattg ttctagttag 69780
ggatttggtt gctttaaaaa atattgaagc taaagggtttt tttggaagaa tatggtcaaa 69840
gttcggttctt tggatatatta gcttatttgg tctaacagat gaaaacacca cttaaaggcg 69900
tttttaaatgg ggctttttgat actgtagata acataaaaaat atcacctttt tctcgagctt 69960
atactttttc tgatagcgta tatgaagtag ttccttttctt taattcaact gcaattgctt 70020
ttgatgatca tataaaaaga ttagaatttt ctgctagtca actagcgatg gccgtagatt 70080
taaaagaagt tatatttgaa attaatctt taataacatc atctgaattt agtaatggct 70140
atgtctatta tcaagttact cgaggcgttg atccattaaag gtctcatatt catgagccaa 70200
atttaaaat agaaaacctt ggctatgcaa aggcgcactt attccaatgg aaaccattaa 70260
gagtatcagt atgtgatgat attaggtggg ggaggtgtga tattaaatca acatcattgc 70320
ttggtaatgt tatgaacatg aatgcagcaa aactagataa ttgtgatgaa gttattatgc 70380
acaaagataa tttgttgacc gaggtggcg catctaact tttttttgta aaaaatgatt 70440
caatctgcac ccagctcta aatggaaata tacttcccg cataaccaga gcattactaa 70500
ttaatgaatt aaaaaattat agtatagagg tcatcgaaga taacttcagg ctagaagatt 70560

tatcaaatgc ttcttgcgcg tggctcacaagcgcaacgaa agggctggcc cccatttctg 70620
aaattagtaa cctagaatct cacttggatt tataccatcc tctttcaaa aagagtgaag 70680
aaatttttaa taagaagttc ctagttaga agtcagttta ataactatat tatcaagctc 70740
gtaccaaaaa tctttatcac taagaccttt atttgcaaga tcaagctcaa atacattttt 70800
ttggagctgt attaaagttt gtagtttatg atttttaata aaatgcatat attcagaagt 70860
tttatttttc caaaccocag atttaattag actatctaca ggattttttac ttgtcttagc 70920
attggctgca ctgttaataa ttttaccac tatccaaact agaagaggtg cataatgac 70980
ttccgctgat tttattgaat ggataatttt taaagcttgc ttggtattat ttgaaaataat 71040
tttatcttca agctcaaatg gcataaattc tgcagactcc acagaagggtt ttcttttcttg 71100
gccatcacca tctttataag tcagtcttag aagcttttact tcatttctgtt gggcaactaa 71160
gtttccagag ttcatttcag atatattttg cactaaagag ggcctatctt tttcagaaat 71220
aaaagatagt tgatgcttaa gccaaatctt ttcttcaaaa gatttaagtt ttccacaatc 71280
aatgatgagg gctatctcat ccactctgtt taccactta gttgtgaaat taagcttttc 71340
attgcttgaa ttaataatta ttgcaatatt atccatatgg ttaatatatt caatttgaga 71400

tattttaatt atttggtcag gtatttttcc ttgatcatga ttaatctcaa ctattatttt 71460
tgagccaaat aaagatccag aagcattttc aataatagtc tgatctattt tatcaaaccc 71520
atctttatta ataattgttt tttctttgaa cccttttatt gataaatgtt tgagcaaaag 71580
atcttttagag ttatttttta agacaacctc agatccaaaa ataaagaaga tattttgtga 71640
ctgatctaaa tattttttgag ttgtcattgc ctacatttc aaaagcatgc acctcaaaaa 71700
ttatttgatc aattaattcc ttctgcattt cagactcaag agatttaatc atttgatcct 71760
gtgcaaatgg attaagttca ttcgatttat atctcttgta gctaacaagc tttttactta 71820
tttgtttatt tccaatcatt ttatatgta tttctaaaatt aatttctcct tcaagagctc 71880
ttagagaaga gcctccatag atatcatatt tattttaaatt atagtcttgt attagaatct 71940
gattttgacg attctcagct tctttttctg tattgaagag gattgctaac ttttgctcaa 72000
aagaattagg gacactacta tcaaaccgaa aattaaatag atctttttta tctagatcta 72060
tatttagttg attaaattgg cagccagtta aagatatcaa caagcaagt tataaaaaa 72120
aaggctcttag tttcataacc tcatagtatc tatctgaaga atatatatca aagatttaatt 72180
ttaaataaca aaatttatta tttttcttt aatataaata gtttttttaa tagccgagtt 72240

gtttattgtt gatgccacat tgtcgatagc cagtgcataa gcttctatat ctttttgctc 72300
aagggttttg tctatcatta ccttacctct cacctttcca ttacttgaa ctattaaatt 72360
aaattcagaa acttctaata actcctctct aagaacaggc caagaagatt caatttcttc 72420
ttgtgcaag tcaaagtaaa aattattcca taaatgttgt gagatatgag gcgcaatagg 72480
attgagagtt ttaagaataa ttattattgc ttcatthaaga caatattgat ttgagatcga 72540
cgcattacct tctttaaaag aatcaggat aaagttaact agtccatta tcgaagcaat 72600
ggctgtatta aatgaatata ttgtttcaaa gtcatattgta accttttaa gattattatg 72660
cgattttctt cggagctcta cttcttggtt cgaaggatcc ttgggctcct ctaagtcaat 72720
aaattttcta ttgctgacta agttccatat tttttcata aatcttgagg caccttctac 72780
agatgactct gaccattcta gactttgctc agggggagct gtaaacatca tgtaaaagcct 72840
tacagtgtca gcgccatact tttctatata agattgagga tctacagtat ttccttttga 72900
tttggacatt ttagcccccatt ctttaagaac catgcccttggt gtaagaagct ttttaaaagg 72960
ctcattgcct tcaactaaat ccatactctt aagcgcctta tgaaaaaatc ttgaatatag 73020
taagtgcagt atcgcatgct caataacctc tatatataa tcaacaggca gccaatattt 73080

tgaattttta tcaaacattt catctgcatt gtcagatgaa gtaaatcttg cgtaataacca 73140
tgatgagtcc ataatgtat caaaagtatc tgtctctctt ttagtcgat cagaaatatt 73200
ataaaaatct tcattttgac ttaagggaaat aggtgccgag ttctttttta gctctggaag 73260
cactataggc atatcctttt catcaataac tcttggttca ccattttcat aaacaaccgg 73320
aatagggcaa cccaatatc tttgtctgct cacaccccag tctcttaatc tgaactggat 73380
taactgctca ccaaggtttg cgtctgctaa atctttaata attcaagag atgcctcatc 73440
agagtccatg ccatcatatt taccagagtt tattaatttg cctttttgta caactggaag 73500
ttcattatta ttatcaaaac taataacttg aggaatctct agattatatatt tagatgcaaa 73560
ttcgaaaatct ctttgatcat gagcaggaac acccatcact actccagttc cataatcaag 73620
aagtacaaaa ttccaatcc atacgggtat cttcttcttt gtaattggat gaattacatg 73680
cattccacta ttaattccaa gtttttccagc tttagccata tcagcttcag cagctttcac 73740
ttctttgcat ttgttttagaa agtctttttat cgattcatta ttttttgaca accctattga 73800
tatagagtga tttgggtgata ttgctaataaaa agaaacacca aaaattgtat caggtcgtgt 73860
tgaaaaatacc tttaaagaat catcggaatc ctcaatcatg tatttaattt ctgctccctg 73920

agatttccca atccaattcc tttagcattgt tttaacattt tcaggccaat ccacctcatc 73980
caaagatgtt aaaagtccct cagcgtagtc tgtaattttt ataaaccatt gatcaatttc 74040
ttttatttca acttgcgccc cagacctcca tcctttttcca tctataaactt gctcatttgc 74100
taaaactgtc tcatcaactg gatcccagtt cactaatgat tttttcctat aaaccaaac 74160
tttatcataa aactttttta atattaactg ctccactta tagtatccg gttcacatgt 74220
tctaagctct tttagaccaat cataccaag gcccaaagat ataagttgct gtttcatatg 74280
ttcaatatatt tgattcgtcc aatcttttgg gctaaacttta ttagcaatcg ctgcgttttc 74340
agcaggaagc ccaaatgcat ccagcccat aggctgaaaa acattaaagt tattcattct 74400
tttatatcta gatatgacat caccgatcgt atagttttctt acatggccca tatgtaattt 74460
acctgatgga tatggaaaca ttgataggca ataaaatttt tctctaccat caggattagc 74520
tttaaatgtg tcttctttta accattcatc ttgaatgggtt ttctctattt gacttgggtt 74580
atattcagga ttcatTTTTT taggaactca ttttctaagt aattaatagt atgtttattt 74640
tctttaagg tatcatcatg cagaatccctt tgatgaagag ggtgatattgt tttaatgccc 74700
tctataaaaa attcatctaa agcactcagc attctcttaa tagccgagct tctagaattt 74760

gcctgagtaa	tgatTTTTGc	tagaagtGaa	tcatagttag	ggggaactct	atatcctcca	74820
taaatatgtg	agtcatatct	tatGCCAAA	cgcCaggTg	tatgcatttt	tgtaattgtt	74880
cctggggatg	gttgaaaatt	atcaggatct	tctgcattaa	ttctgcattc	tagtgagtgg	74940
ccatgaaaat	taatatcatc	ttgattttagc	tcaattggca	tttcaagagc	aattcttagt	75000
tgtgctttta	ctaaatcaaa	gccagttatc	atttctgtta	cggatgttc	aacctgaatt	75060
cttggtattca	tttctataaa	atagaattga	ttatcttctgt	ataaaaattc	aatagtgccca	75120
accccttcgt	aatttatTtc	ttcacataaa	ttgacacaag	ctattagtgt	tttattaaaga	75180
gcttcttgat	ttataattaag	tgctggagct	tcctcaatta	tttttttgatg	tctcctctgc	75240
atactgcaat	cccttgTacc	taaatgtatc	gcctttcctt	taccatcacc	aactattttga	75300
acttcaatat	gtctagggtt	cccaataaat	ttttcttagat	aaatgggtttc	attgccgaat	75360
ccatttttag	cttcctgcat	tgttattttct	gcatgtccaa	taagatcttc	ctctctttca	75420
acaactctca	ttcctctttcc	accgccacc	gccgttgccct	taatcataat	tgtagagccg	75480
atatctctag	caattttttt	aaactcatca	ttatcagatg	gaatctcatc	tttataaacct	75540
ggaacgattt	gaattccaga	tttttcagcc	agtgtttttg	ctgttatttt	gtcgccccatt	75600

ttttggattg tttctgaggt tggccaata aatttaaata cacttttttc acacatttct 75660
gcaaaattat gatcttctgc aaggaacca tatccaggat atatagcatt agcacttgta 75720
agttctgcag ctgagagaaat ggcaggtata ttaggttagc tttgagttgg agatgcaggc 75780
ccaatacaa cagtttcatc agagaatctt aaatgtttta gatctttatc agctttctgag 75840
taaacagata ctgtcttgat accaagctct ttacaagctc ttatagctcg aagagcaatc 75900
tcaccacgat tggcaataag aactttataa ctcattttaa ttgacagtta taattttttg 75960
tccaaatca accggacttc catcttcac gtcaatactt gaaatcttgc catcaaatc 76020
tgatttaatt tcattcatca ttttcatggc ttcaactata cagagcacat cacctacttt 76080
tatattgtct cctactttta caaatggatc ttttctggga cttgggcttc tataaaaagt 76140
tcctactatt ggagaggtaa caacatcccc tatgacagtt tgcttagcct catccaattg 76200
atgagttggt ggtgagatag ctggctcatt tttgacaatg attgggttttg aaacaaatc 76260
ttgattagag ccgttatctc ttgaaattct tactgactca tctccttgac ttacttctat 76320
ttctttttaa tcagactctt gaagcatttc tataagtttt tttattttcc taatatccat 76380
ttatttacc ctgtatttta atatgccctt aagcatgcc tctcataac ccttagcacc 76440

aataccagtt atcactttctt cagctatatc agatagatat gagacatgtc taaattcctc 76500
tcttgtatat atgtctgaaa ggtgaacctc atagaatggt atgttgacac caagaaatgc 76560
atcccttata gctatgctcg tatgagtata tgcagcagga tttattatta caaagtcttat 76620
tttttcttca atagtgtctat gtattgattc tattatttca tgctccgcat tacttttgaaa 76680
agataattaaa ttgcattgat tttgatttgc taagactaat aaatccttgtt gaatatcttc 76740
aagtgatttg ctgccatata cttccttttc tctagtccg agaagattga ggttggggcc 76800
atttataagc aatatatttca tataatagat tttaacagaa atttacagat tttttaggat 76860
tttaaagata tttttataag agttgatatt tttgcatggg atagcaaaaa cccgcttcag 76920
agcatccctg atagtatat aaaacctccg ataagtcaa tgaagaattt atttcaatgc 76980
ttataataaa ttcattctcta agaattcag tctcgccaaa aaattcatcc ctatatagat 77040
ctttactga ctcaagcgtc ttaaatggga gcactttatc tttaaatttg agctctatac 77100
tatccaagta catataataa ccatctctta ttggccaagt tataatgct gcttgttttt 77160
caacattggt tgttaaagca aatacttcat tggccttctgg gaccctattt gagtttttcaa 77220
aaagatttgc tgaattcaat tcacctgcaa tcacattgga tagtagtata attattaac 77280

ttttgatcat gaagacttaa gaataaaccc ctaagtactt aaaataaaga ttttattaaa 77340
taacacggta gttaaaaatt tatgttccag actttcgac taaaaatttt cttttatatg 77400
cctatttggc ttttaaaat aattttttat agaaaaagaa cagtaataag agggcatcaa 77460
tttgatgctc aatctgctgc attattgtct ttgcttccaa agaaaagattt atctgaatta 77520
ttagatggtg aaattgcaa agctagaatt actcttgaag aagcaagaat tcaaaataaa 77580
gtttctttaa caccgtctat acaagtcaga aaagtagatc atattttacc aaagcatgat 77640
ttaattctca gagaatacaa gccgcatcaa gaggatttaa aaaaagttaa tctctatttt 77700
catggaggag gttatgtcct taactcagtc aatacacatg atgatatggt ttcatatatg 77760
tcagataaaa taggagtaaa gttttattct ctagactata gactatcgcc tgagagcaaa 77820
tatacctgatt ctcttgatga tgctctagat gctttttctt ggctcatcgg ccaggatat 77880
ggaccaggtg atatttcagt ttgtggagac agcgaggcg ctcatctagc tgcttctttg 77940
tgccattacc tagctgaaca aaataaagat ttgcctagta gccagctact aatataatcca 78000
atgtgtgac cgtccctgttc atctgagtct tataatttat tatcttcagg atatcttcta 78060
actaaaaaa ctatgatttg gttttgggat aagcttagaa cttccgaagt aaacaataac 78120

gattcggcct ttaatctctt aaaatttaaat tttagaaaaa cttaccgcc aacaattatc 78180
gttacagggtg gctttgatcc tctatgcgat gatggagaaa aatatgcata ttattacat 78240
aaaggtaaac ataatgtgaa acaattacat tatccaacaa tgtttcatgg gtttgcata 78300
atgactaaat tgaagcagc gcagatagcc gttgaagatt ttttaaaaga atataagaaa 78360
atactatgag taaaatttta gaagtaagt acttaagcat caatttttca acaagggatg 78420
gattgtttta tgcagtggat aatataagtt ttgatataga aaaaaatcaa accttggcct 78480
tggttggtga gtcaggttct ggtaagtcgg taactgctat gtcaattctt cagctccttc 78540
aaaaaccaca agcatcatat tccaaagagt cttctattaa gtttaatggc gatgagataa 78600
taaatgccaa gtatgaaaag ttactttcct tgagaggaaa tattatatct atgatatctt 78660
aagagccgat gacctcacta aacccttct acagagtagg taatcagata actgaatcaa 78720
tactacttca ctcaaaaagc tcaaaaaaag atgcaataga tgaagcaaaa aaattaatgg 78780
cacttggtga gattgatgat gttgaaagac ggttctatgc ataccctcat gagctttctg 78840
gagggcagcg acaaagagtt atgattgcta tggcccttgt taataaacct gagctattga 78900
ttgctgacga gccacaacc gctcttgatg taactatcca agcccagata ttagatctca 78960

tgtctaagct taaaatgaa ctaggcatgt caatactttt tattactcat gatctaggcc 79020
tcgttcaaga attttctgac aatgtttgtg ttatgaagaa tggcaagata gttgaacaag 79080
gaaataactgt tgaggatatt aacaaccctt ctacagaaata tacaaaaaaa cttttagatg 79140
cagaacctca gccaaaatta gataaccctg taagtgatga gccaaataatt gagatcaatg 79200
acttaaatgt ttactactca ataccttcaa ctaatttttt taaaaagaat acttttcatg 79260
ctgttaaaaa tacttccttt agtatattata aaaatactac aattggcctg gttgggggaat 79320
ctggatctgg aaagtcaact ttgggtaagg ctatagcaaa cttagtttct tataagggga 79380
atattaaatt tgaggggaaga gatatcaact caaattccca aaaagaaaaat aaagaattga 79440
aaaaaaatgt ccaaattggt tttcaagatc cttatgggtc attatcacca cgaatgacag 79500
tgggggagat agttggtgaa ggtttaggtg ttcaactttaa gcttacaaaa aaagaaaagag 79560
acgaaaaggat agataagggt ctgtcagatg tcggtatcga aatagtagct aagaataaat 79620
atccgcatga gtttttctgga ggccaaaaggc agagaattgc aattgctaga tctttaataa 79680
tgaatccctgc ttttatgatt cttgatgagc caacatcagc attagatagg tctattcaaa 79740
ttcagggaat cgattttattg aaagagatac aaaatgaata tgggcttact tatcttttta 79800

taagtcatga tttaaagggtt attagatcga tgtcagactt tatttttggt atgaaaaaatg 79860
gagaaatcgt agagtcagga cttctcaca aggtcttga aagcccagag caagactata 79920
ctaaaaaatt actatcagct gctttaaagt atgcattctga ataattaaat atatymcmtm 79980
tggcaaatag aartattca aaagagctcg ttgacgggtcc taatcaagct gcttcttagat 80040
caatgcttag aggagtaggt ttccatctcg aggmthttcac aaaaccattt gttgggattg 80100
cttccacagg agcaaaaagta accccatgca atatgcacat aaatgcactt tcagagatcg 80160
ttgagaaatc agttgtagt tcaggaggaa aggtgttct ttttaatact attactgttt 80220
ccgatggaat ttctatgggt acacagggtg tgaatatattc tctgtttct cgagaggtaa 80280
ttgcagattc aatagaaact gttgtgggat gtcttggtta tgatggagtt attgctgtcg 80340
gtggttgtga taaaaatatg cctggatgca ttattggaat ggcaagatta aatagaccat 80400
caatatatat atatggtggt tctatcaaac ctagttaaaga aaataccgac tatgttactg 80460
tttgtgagaa aactggagag tactcaaaaag gcgatcttaa agaattctgaa ttaattcatg 80520
tagaaaaaat ttccgtaaaaa gggcctggat cttgtggggg aatgtatacg gcaataacta 80580
tggcttctgc gattgaagct ctaggcatga gtcttcctgg aagcagcagt caagatgcaa 80640

tttcacacga caaagaagat gattgtttta aggctggcga agcgataatg aatttattag 80700
aaaaagatat taagccttca gatattatga ctaagaatgc ttttgagaat gctataacaa 80760
tggttaattgc tctaggaggt tcaactaatg cagttctgca tttattggcc atggcgcat 80820
caataggggt tgatttagag ctatgatgact ttacaagaat aggaaaaaaa acaccgtta 80880
tggcagatct taagcccttt ggttctcatt atatgtctga actcaatgct aatggcggt 80940
ttcagccact aatgaaaact ttgcttgaga aggattact acacggcaat tgtcttaccg 81000
ttaccgggtca gacgcttgct gaaaatcttt ctggaataaa accttacgag cctgatcaag 81060
agataattaa atcatttgat aatccaatta aatcaaatag ccatcttaga attctgtatg 81120
gcaacttagc gaaagatggt gcagttgcaa aaattacggg taaagaaggga acttcctttg 81180
aaggaagtgc tcgtgtat 81240
ctataaaagc tggagatggt gttgtaatta gatatgaagg gccaaaaagg ggtccaggca 81300
tgagagaaat gctaaaacca acatctgccca taatgggtca aggtcttgcc gatcaggtag 81360
cttttataac agacgggtcgt ttttcaggag gcactcatgg attgttgtt gggcacatta 81420
ctccagaggc tgcagatggc ggcttaattg caataattaa agatggcgac tctatatata 81480

ttgatgcaga ttctgataag ttaattctta atatttctga ggatgaaatt tcaaatagac 81540
taagtagatg ggtaaaccca aaaacgcctc ccaaaaaagg agtcttagca aaatttgcaa 81600
aaagtgttaa atcagctagt ctggagcgg taacagatta aatatgtatt ttaaaagaaa 81660
atttccaat agtagactaa gaagaatgcg tctcaattca aacctcagag acttgcttgc 81720
tgaagttagg ttgtctacaa atgatcta atgcaccatta ttataaaag aaggcttgag 81780
tgaaaaagag gctattgaaa gcatgcccaa tattaataga tatggacaag attcaatttt 81840
ttcagaaaata gaagagctac tagagcatga tataaatagc attgccctat ttccagttat 81900
taatgattcc aagaaaaaca gcactggaga tgaggctatc aatgcatcta atttgatgtg 81960
tgaaactatc agtaacataa agaaacgatt ccctgaaata atcttaatat ctgatgttgc 82020
tctagaccca tataccgac atgggcatga cgggtatttta aaaaatgatt atgttgataa 82080
tgatgagact ttagctgttt taaggaagca atcactaaca ttagctcagg ctgggacaga 82140
cataattgca ccatcagata tgatggatgg aagaatagc tctataagag aggcattgga 82200
cgaaatggat tataaaaaata caatcctact atcatatgca gcaaaatata attcaaaagt 82260
ttatgggtcca ttagagatg ctgttaattc agcttcaagt ctaggcaagt cttcaaaatc 82320

cagttatcaa atgtcaccaa aaaatataaa tgaagctcta catgaagttg ctatggatat 82380
taatgaaggt gcagacatag tcatggttaa gccaggtag cttatctag atataatttc 82440
aaaagtaaaa gaaaccttca aagtacctac cttgcatac caggttagtg gtgaatatag 82500
catgcttaaa ctggcgattg ataaaggatg gcttgaagc gatgttatgt tagaatcatt 82560
aataagtttt aaaagagcgg gagctgatgc aattctaacg tacgcagcta aagaaatttc 82620
caaggagata actaacaatat gagcaatgtg atagaaattc gtgatgaaga aagctttaat 82680
agcgacgtct taaattcaga aaaacctgta ttggttgatt ttgggctga gtgggttgga 82740
ccttgtaaac agcttgcacc aactgttgaa acagtgcag cagaaaaatc agaaacatta 82800
aaggtttgca aaatggatgt tgattcaaat agagagattg ctgctaaata tggaataagg 82860
tcaatccctt cattaattat atttaaaac ggagagcctg caggagttga agtagtgct 82920
ctaaccaaac aacaattaga ggactttata agtacagtag tttaactttg caaagacttc 82980
tttgcataat gcaagaaata ggattatcat ttcgacttct aaggctgaaa gccactcaaa 83040
acaaaaaacc accttttcaa tcataaaaaac aacaactaga acggaataaa ctaaatgaac 83100
cttactgaaa ttaaaataaa accaataaat gaactttag atatagtac tgagcttggg 83160

cttgaggatg ttggaaggct gaaaaagcaa gagataatat ttagaataatt taagcataag 83220
gcttctgaag gtgttgatat ctatggtggc ggagttcttg agattttaaa tgatggtttt 83280
gggtttttgc gatccccaga aggctcttat tgcgctggcg aagatgatat ctatgtttca 83340
ccaagccaaa taagaaagt tagcctcagg aaggagact cagttgctgg gaagataagg 83400
accctaaag ataaagagcg ctattttgca ttaatccaag ttgatactat taatggtgaa 83460
gagccaagaa agactaaaa caagattctt ttgaaaaatt taactcctct ttttcccaat 83520
gaaagactaa tccttgaaca aggaacaggg tctaatagaag atctttcatc tcgaataatt 83580
gatttgattg ctccaatagg aaaaggtcag cgtggactta tcgtttctcc acccaaggct 83640
ggtaaaacct taatgcttca aagcatagct cattctatta aaagcaataa tccagaagta 83700
gagcttatag ttcttttgat tgatgaaaga cctgaagagg taacagagat gtcaagaact 83760
gtaaaaggag aggtttagc tagtacttt gacgagccac ctactcgaca cgttcaagta 83820
gcaaatatgg ttattgaaa agcaaaaaa cttgtagaac ataagaaaaga tgtagttatc 83880
ttattagatt ctattactcg tcttggaaga gcataataact cagttcagcc tgcatacagga 83940
aagatataga gtggtggagt tgactccaat gctcttgaaa ggccaaaaag gttttttggt 84000

gctgctagaa atcttgaaga ggggtggaagt ctactattc ttgctactgc tctagttgaa 84060
acaggctcaa agatggatga agttatttat gaggaattca aaggtaaggg taatatggag 84120
attcaccttg aaagaaaaat agccgaaaaa agaatatatc ctgctattaa tattagaaga 84180
tcgggaacaa gaagagaaga ttactttact gctgaagatg aattacaaaag aatgtgggtc 84240
ttaaggaaaaa ttttagacga tatggaagat gctcagtcaa ttcagttcct aatagataga 84300
ttgaaatctc ataaaacaaa cgatgagttc tttacttcaa tgaaaggggg taatggcaag 84360
aagactagat aaagtttttt gccatatcaa tcatcatctt atcagttggg ctttgcgata 84420
cgtgaatttc taattcttta aactcattt gacatctatc ctttataatt tctgaagcaa 84480
caataaatat tttttttctt aaactgcat cgtccagatt tgttattaag aatttaagag 84540
tactaaaatt atagattaaa aaaatttcat tattatctgt aacttttggg atttgctcca 84600
gcaaatagat tacttcatag caaactatct catctagact agccttaagt ttttgttgaa 84660
gaaaaccatt tgaattttcg ccgcaaaaata aaagactctt tcctagaaaa tttttctcta 84720
ttaacttaag aattccttca gatgaatggc tttgtggaaa atgagatttt atgccacttt 84780
caagaagttt attggaagtt gcgggccccaa cagacaaaaa attgtgtgga agatcatcta 84840

aatcgaaaa agatttttaa atatcaagtc cgtacgaagc agcagcttgg ctagtaaaaa 84900
ttaagtttga atatgaatga atattttaa tttatcgat agcgattttt gaagggtcga 84960
ttgaattaat tttagaaaaga taaatatattt taagagctat ttcttgagct tcacataaag 85020
atattaggtt tcctgataaa tttttgggcc tcgtattaat tatcattttt taaaatttaa 85080
gatttcgctc ctccaaaat aaactcatca gcaataaaca taatatctat tgcaatatct 85140
tgaaagggaag atatcttttc ttataaaatt ttttcgttgc cttcataaga taatatattc 85200
cctgatatat ttatttggcc attttatct tcacataaaa tagctattgg agacaggcaa 85260
gtacctcca tggcagcaac aaaagatgc tctgcactgg caagaattaa ttcttttaggg 85320
tcaccaatat tttctaaaag ctcaataata tcttttttat ctgacagaca ttcaatttgc 85380
atataccct gagatgcaga aggtaacatt tcttcaattg agaattcata tgaattttgt 85440
tttaaaccac gcctttttat agcagccctta gcaactacca gcccatcaaa cagaccatca 85500
ttgagtttct ttatttctagt agctatgttg cctcttactg gaatagtttt tatatcaggt 85560
cttaagttat taatttgtgc ttttcttctt ggaccagaag taccaatagt tgaatttcta 85620
gcaaatcag aaaatgattt cccatcctta gatagaagca ggtcttcttg agattcccta 85680

ctaaaaac ttattattt aaactttgga tcaagctttg ctggaacatc ttttaaaacta 85740
tgactgcaa tatctgcttc atcagcctct agagaggact caagtgtaga aatgaatagg 85800
ccctttccac caatttcacg aagaggttta tctgttttgt ctccttctga tgtcatagga 85860
acaagctcaa ccttaatat attgatctta gctaataact catctgcaac aaatttttgc 85920
tgatacattg ctagtctga ttgccttgta gctattctta ttttcatttt gttccaata 85980
aaagagcctt cacatcgcca aactctttt cttcaagat agtgaattcc tcaggaattt 86040
ccaaggtagt aaatttacta tactcaaggt agattttaca agatggcttt atttcattct 86100
ttcgaattat agattttaga acttttaatt catattcttc gccaaaagga ggatcaagaa 86160
gaattaaatc aaacgagat aagtcatttt ttttaatcca actaaatgca tccttaaaaa 86220
aaactttaga tttatccttg atgccaaagaa gttgaatat ttttgctaac acagaatagt 86280
tttttttatt aagttcaacg aacacaaactt tttttgactg tctggatat gcttcaatac 86340
caagcgacc cgttcggga aacaaatcaa gacatatata attttcaatc tcaaatgaa 86400
gccaatataa aattatttct ttttaatttat ttgaagtgg tcttaaggaa tccttaaatt 86460
cgaaaggat tttttacct tttaataaac ctccagtgat tctgatatgg tttttcattt 86520

tttttgcaaa atgggtatgc ctatttatta aataaatata attamaaata catttaaaact 86580
aagttataat tcatccatgm gtcctacaag ataatttttag aaaagcaatg aggagctata 86640
tttactctgt cagtgtgatg tcaaatgttg acgagaataa aaagttagt gctataacag 86700
tttcttcagt tacctctggt tctttagatc ctcctagttt gctcgtttgt atcaataaat 86760
cagctggaat tcacaactca ataaaagaag ggtcctctctt ttgcataaac cttttaaata 86820
aaaatcaaga agatatattt aatctatgca gttcatttaa gaccgaaggt gatagattta 86880
atagtgtga ctgggattta agcggcactc cgtttttaaa aagtgccag gctaataattt 86940
tttgtactgt tgatcaatta attcatacc acacacatac tattgtgatt ggtcatgtca 87000
caaactctct tagcgatgag aaaattaata cactgacata tgttgatggt agctatgaat 87060
aaattttcaa aaaatgtatt ttttatttta ataatactca actctttttt tcttgcctct 87120
aacatttttg ctagccagga agagtgtgaa gaaaagccaa gtgttttttat tatctctcct 87180
caagacgggt ttatctcaga atctaataat gtaaaagtct tatttggatc aaaaaatatt 87240
gaaataaata cagctggcaa aggtgagatt gcaaaaaata aatgttttgc aagcgggcat 87300
caccatcttt tagttaatat cgaagcatg ccagagagct ttattccttt tgacaagggt 87360

tatttacatt ttggaggagg tcaaaactgaa acaattcttg atcttgatcc tggaacctat 87420
tctctccaac taattcttgg atcttatgtg cataattcaa aaatgcaggt aaataacttt 87480
aaaggtcaag gacccttttt atcagaaaaa ataacaatta cagttaatta gagattagac 87540
cagttataac tttatctaga tgttcatcta agttattagc aattatattt gcactagaaa 87600
taggaccaga attatcgtca taatcatcac caactacatt cacctttctt attatccaa 87660
caagaccatg caacatagac caaagagtta tacatttaaa ggcaatcact toctcaggct 87720
cgtctgctag attagcaaaa ctttttctca tattatcgta tgttccattt gcagatttaa 87780
gaagatctgg gtaatcggca aagttaccaa cagctgttcc aaacatcaaa tcatatgtat 87840
gtgcattttt taaaccaaac cctatatatt tgcttgcatt agtaacaagt tttttctttg 87900
taattttttt tggattctca aaaaaaacia gctcatttaag tttttttaaag ccaacgggtg 87960
caacagcggc atatacacat tcttttggtt caaaatgcct ataaggagca gtttgagaaa 88020
caccactttc ctttgcaagg gatctaatac ttagtttagt gtaaccatct ctatcacaaa 88080
gcctgcatgc gcattctata agttcttctt ttaagtttcc gtgatgatag ttattcataa 88140
tttaatttta atgttgacac tgctaacata tatcatgtta ttatgtatac accgcataca 88200

ttaagttgat tacatattaa agtcaaatac agatatatga acataacaaa aataaataac 88260
accgtcatag ctctttttatt aggcaatgcg tttctatcaa atcttgaggc tttagaagttt 88320
cttgaggta aaatgcttga tgaatatgct gtaactagag aatttcctgg aaagctcatt 88380
ccatcagacc agtctaagct agcatttgaa atacctggaa agataaactc tattaatgtt 88440
gatatcggag atgagggtcat cttaggggat gaacttgcct cattggatga tagagaagct 88500
ttagcacaaat taaatcaatc aaagggtctaaa tttgatttag ctgaacaagt actagcaaga 88560
tatataaatc tcagagcaga tgggcataatt tctattcaag atcttgataa ggctgaatct 88620
gatcaaatag tagctaagtc gcagtatgat ttttatagag ttaaatttga gcaaaactaag 88680
ttactagctc cttttaatgg agttattcaa aatagatttc tagatacagg atcagtaatc 88740
aatgcaggtg tccaagtttt agaaatttta ggctctagca atgttgaaagc aagaatttct 88800
attccaatga actatatgaa caagggttaag attggagatg agtatgaaat tgatatcaga 88860
ggaatatcta caaaagctac gttagagaga ctgggtctcca tgtcaaccgg aggctccaat 88920
aatagggttag caatttttttag atttgatacc tttttttaatc caggatcaat agtaaagctc 88980
aaactaagca tctctgagaa agcaaaaaggt acatgggttc caattaagtc actgtcccag 89040

tccgaacaag gtatttgggc catataacc attaacgagc aacaagtagt tgtagagat 89100
cttggtgatg ttatttattt tgaagacgaa tatgcttttg tcagcggaac acttaataat 89160
ggtgatttag taattttagg cggagctcaa aaaattattg aaggaaaaat aataaaaaaa 89220
aatgaatgtt attaattttc taatagaaaa gcctaggata ttattttctaa ctttagcatt 89280
tatattactt tctggaattt cttctgggct ttcagttcca attcaagaaa accctgaact 89340
ggctgagaga tggggagggtg ttctgtattt tcttcccggt gcattccccg aaagaattga 89400
aacagagata gtaaatgac tagaaatcaa acttagagaa gttgaagaaa tcgatgagct 89460
tgaatcaatt attactcaag gtttttcaac aattgtagtt gaattaaatc aaagtgtacc 89520
tcctatactt attgaagaga cttgggtccaa ggttcaagac aagctcaatc aaatagttat 89580
tcctcaaggc gcagaaatat ttcttgatag aagcagtggc ccgcctatca ctgttcagta 89640
tgctgtaacc tggaacggca gtggagatgc tccactaata atgatgtcca gactagcaag 89700
ccagctaaaa agaaaaattaa gctcaatagg ctcatctcat caaactgcaa ttttgggtga 89760
aacagatgaa gaaattttta ttgaactaga ttcatcaaag ctatcttcgc ttggattatc 89820
atttcaagat atcgcaagtg ctattcaagc cctagatgca aaaaaacctt ttgggtgtatc 89880

ctcaaaac aattctgagc ttttatatag actcaaagat aatatacaga gcattcaaaa 89940
actctcagaa atacctatca aggttatttaa taaatcagag atcatacagc tagatgatgt 90000
ggcattttatt tcaaaaaatcc cggtttctcct tattgaagac atattcttgt ttaatggaaa 90060
tgtagttatt tctgttgctg gaaccggatc attttctcaa agagtccatg attatgtaga 90120
acgcgcaaca attgttgtag atgagatgag agaaactctg ccgactgaga tcactataga 90180
tttagtttat gacgaatctg cttacacaac taaaaaattt aatgagcttg taaaaagttt 90240
ttcattagca atattttttg ttttagcttt aagtcttttt tttcttgga ttagatcagc 90300
aataattgtt actcttatcc tgccattttc tatttgcctg gttatgattg gttgtaggtt 90360
tataggctta cccttgcata tgacatctat cactggaatt attattgcac taggattgct 90420
tatagataat gggattattg ttgttgaaga ttataagaat agaagagcat ctggattaaa 90480
tatcaatgat tcaatttcac aaggactaaa aaacttatgg gctccattag ctgctgcaac 90540
agcaacaacc gtcttctctt ttcttcctat tgttactgga gaagggtcga gtattgaatt 90600
cgtaggcggc atggcaatga cagtcattat gtctataaca tcttcattag ttttggcggtt 90660
attaatgggt ccagttctga tgagttatat ggaaaaaatt ccgttcttta aggatgtgga 90720

tattagcaag gaaggggtata gaaatgaaaa aatccttaat aaatataggg ccttttttaa 90780
ctgggcgttc ttagttccta gaagagcaat catgatatcg cttgcattgc ctgttctagg 90840
attctttctt tttaattctt tacctaaaaga tttcttttctt gctcaagata gagatatgtt 90900
tagagttaat atagaactgc cttctaacgc ctcatcactt actacaatgc agagagttaa 90960
ggaaattaga gaagatatct tagatagtga tttaatttca atagaaaaag attattcgtt 91020
tatcggcaga atgatgccta gagttttgat gaatgttgtt ggtggagaag aaaaaaagg 91080
atccaataat attgcgcagt ctgtattttt tgctactgat tattatgaaa tgattgaaaa 91140
ccttcagat ttatcaagaa gactgggttaa aaataaccct gacattatag ttaatataga 91200
tagtttctcg tctggccccc cggttttttc agatgttagt tatgtaattt ttggagatga 91260
tccagattta ctaaaatcac ttggtgagga gctagagcta attattaaca attctcctga 91320
tgtgagtctt acgaaatctg caacttcaaa ctcaataacc aatgttgagt ttgaacttaa 91380
cagctcaaat atttcaactgt ctggtcaaaa tgccaattat cttgtaaatg aaatgtttac 91440
tgcaaaacaat ggaatatattg ttggcactat gttggattca acaaaaagaa taccagtcag 91500
gctgaaaagg ctgtctaata aaaacaatat tacgggaaat actagtttta taacaatgcc 91560

ctctcaaggt ggttttgagt attttgatag ttctggaaaa agctcactaa caaacaatac 91620
gtcaacaatt actaggcttg atggccaaag acaaatgat gttgagggct ggatttggac 91680
aggtagctc ccgtctgcta ctgaaaaatc tattaataaa gatgttaaa attttgaatt 91740
aagattgcca ataggctatt cattaaaaa acttggcgag gctgaaagca ggggccaatc 91800
tcaagcctca ttatactctt cagcttttat gtatttcatt ctataatag taggcttggt 91860
tatggcgctt aattctttca gagaggctgg tctaattttg tctgttgcac tcttatcaat 91920
tggactatcc tttcttggtt tatttatagg ccagcaaat tatggattta ttggaaactat 91980
aagtgcata gggttaattg gcttatcaat aaatgattca attattgtct tatctcacat 92040
aaaagaagag gctgagaaga aatcactaac caagctgag ctgtttgaag ttgtttatcag 92100
gtctacacgt catataatca ctacctctt gacaacactt ggtgggtttc ttccacttat 92160
ttttgcaagt gtattcttca aaccgcttgc ttgggcaatg agcattggag tattaggcgc 92220
gactattaca gccttattat atattcctgc aatgtttatg ataatgagaa aggttaagta 92280
ctagaacaac ttcccgagca ataatttctt cctagctttt ttacaataag tgattcatga 92340
gtataggctc cgcaagaatc acatttaac atagagtta caatcttttt ttatgaggt 92400

cctgatatga gagatctgaa aaagaatacc gcagttatta tcaatataaa tatgaccaat 92460
ggaatcaaaa gtaatatcga ttttaataag aacatttttt agtcactaga tgctggctta 92520
ctatttctta taaaattcca agaccagcca gacttatctt cagatttttt tgattgctca 92580
tcttggtagt taatagaaat tattttttta gtatcttcaa gcaattcaat atatcccaaa 92640
gattcatatg aagcctctaa gatctttaaga gctctataat tttcactaga gttaggaata 92700
ttttcaatta cataatttgc tcttcttatt gctgctatat gcgcatcaac actaaacgtaa 92760
taatcagctg ccgcaagctc atttcttgca atcatatttc ttaaatagat atttctttgt 92820
ttagcatatg tagaatattg gctatcagga aatcttgta agaattcagt tagttctgaa 92880
aatgattcct tagctcctga gatatctcta tttgatagat cagtatctgt cattcttaca 92940
ataaagctat tatctctcgt atagctagaa aggcctttca taaaatatgc ataatcaata 93000
tttgggatgcc taggatgaag tcttataaat ttttctgctg cagcatgcg agcctcagtt 93060
tcagcattca taaattgagc ataaaataagc tctacttgcg cctgttcagc atatttgcca 93120
aaaggataac ttgattcaat tgcttctaata gaatcaatag caccaaaaata attttttcca 93180
gccatccctc tttgggcttg atcgtaataa attttttcag gctgttctat ttctgggcca 93240

tcagaattac aactaaccaa taacagagtt acaattggca atactataaa tagttttaaa 93300
atagtttat ctttcattat ttgcacattc tacctgcatt atcgtaattt aagcttgaaa 93360
acatctatct agattaattt ttttataata agttcatatg attgttaaaa atgttccaaa 93420
tgatctatct tcaatgaggc tagacaaagc tacagcagaa atgttttacag attattccag 93480
aactcagata aaaaaatgga tagaagaagg cagagttctt ataatggag aggtatccca 93540
gccaaagat aaagtattatg agaatgatca gattgaatta agccctaaa aagaacaaaa 93600
agtatcatgg gaagctcaag atatagattt tgaaattcac tttgaaaatg aagatttttat 93660
tataattaat aaacctgctg gtttaataat gcattccaggt tctggttgct atgatggaac 93720
tctcgcaaat gggctcattt ataaatttcc agaattgac aatatccca ggtcggggaat 93780
tgttcatcga ttagacaaag atacttctgg cattctgctc gtagcaagaa atgagtcttt 93840
taggaacttt tttattaatg aaatgcagga gaggagagt gtataaaaa atacgtctat 93900
tgttattggt tctacactag gaagcttttc tatagaagag ccaattggaa gagataaaaa 93960
taatagaacc aaaatggcaa ttcgagaaga tggcaaatg gctttaacat ttgtaaaagt 94020
taaagaaaat attggaaact actctgtgtt ggatataaga atagagacag gaagaactca 94080

tcagattagg gtccatctat catcaaaaaa actaccaata attggagata aaacctatga 94140
cccaagcagg tctattgcaa gagatacccc tgaagagcta attaataatta tccgagggttt 94200
tccaaggcag gcattacatg caacacacct ctcatccaat gaccaaaaaa caaataatat 94260
ttttttctttt gatattccca ctccaaatga tatggaggaa ctactttctag aattaagaaa 94320
attgatctaa tagtaactaa aaacttgttt ttgattaat aaaataatat aaaccttatt 94380
cctaagagtt ttttggttaag aaattgaaat tatctggcgc agacatgcta atgcaagcac 94440
ttcatgatga aggtgttgag ctaatctttg gctaccagg tggagccgcg cttcatatct 94500
atgatgcaat ttttagacaa gataaaatag atcatatttt agtaaggcat gagcaagggtg 94560
caaccatgc agcagatgga tattcaaggc cgacaggtaa gccaggagtt gtcttagtca 94620
cttctggacc tggtgcaaca aatgctatta caggaatcgc gactgcattt atggattcca 94680
taccaatggt agttatttca gggcagggttg ctagccattt aatagggtact gatgcttttc 94740
aagaaactga tatgattggt gtttcaagac caattgttaa gcatagctat acagttttta 94800
atgctgaaga aatacctaag ataattaaag aagcttttta tgtcgcaact tcaggcgagac 94860
ctggacctgt tgttatagat atcccaaaag acatgacagc tccggataat ctttttgatt 94920

actcgtatcc tgaagaagcc aagataagat catacaatcc tccgattgag ccagaaaaaa 94980
atcaaataga tagagcagtc gaagctatat tgatatcaa aaaaccagtt atatatgctg 95040
gtggtggggc aattgctagt aatgccgaaa aagaattact tgaacttaat gaaattattg 95100
atgctcctgt taaaaatact ttaatgggat tgggtattta cctgctagt catcatagat 95160
ttcttgggat gttagggatg catggaacat atcaggcaaa tatggcaatg cataatgcag 95220
acttaataat tgctattggc gccagatttg atgacaggat taccaataaa ccatcaaatg 95280
ttgcacctaa tgccaaagtg gttcatctag atgttgatca ctcatctgta tcaaaaatta 95340
tagaagcaaa tgtagctgtt tttaggcaag taaaaaattc cttaaaatta ataaaagaaa 95400
ctcttgaaaa aaaattagac tcttacgatt ctctcgctct tcagcccttg cagatcaga 95460
taaaagaatg gaaatcacta catggtttaa attatgagct ttataaagat gaatctgatg 95520
atcatcccat tttaaccagg gctgtagtcc agcatgtcca tgagattaca aatggggaag 95580
catatgtgac ttccgatgtt ggtcagcatc aaatgtttgc tgctcaatat tatcattttg 95640
ataagcctag aagatggatc aattctgggt gtctaggaac tatgggtttt ggtttgccag 95700
cagcaatggg tgtaaaactc gcttttccaa aagatgaggt tgttgcatt actggtgagg 95760

gtagtatcca aatgtgcac caagaattgt ctacatgtct tcaatataat ctccaataa 95820
aaataattaa tatcaataac gaagctcttg gtaggttaa acaatggcag gatatgaatt 95880
atggaggaag gcactctgaa agtacctatc aaaactcgtt accggatttt ataaaaactgg 95940
ctgaatcata tggtcataata ggaattaaaa ttacaaaaaa ttctgattta agtgaaggct 96000
taaaaaaagc ttttgaaatg aaagataaac ttgtctttgt tgataatttat gtagatcctt 96060
cagagcatgt ttatccaatg caggttgcaa atggcagcct agaaaaatatg tggctatcaa 96120
aggatgaaca aacatgatta aaagaaaaact aattttaatt atggaaaaata aaccaggagc 96180
tctagtaaga gtagttggac tgtttcatca aagaggctac aacattgaaa cccttcatgt 96240
agatactgtt aaagactttt ctacttaca atcgatatgt aaaaaaaact tgaaccacaaa 96300
tttgaggata atcaaatatc tagactgacc atagaaacaa tggtttcaga tgaccttatg 96360
aggcagattt tgagacagct caataaatta atagatgtta tagctgtaag caatgaagag 96420
acaacctatt taaaaggagt attattagat gaaaatttat tatgacgaag atgcaaacat 96480
agaaaattatt aaagggatga atgtctcaat aattggctat ggatctcaag ggaatgctca 96540
tgcaataaat cttcatgaat ctgggtgaag tgttactgta ggtttaagag aagggtcttc 96600

ttcatgggca aaggcagaag aagcaggctt aaaagttcaa acagttgctg attcggtaat 96660
ccaagcagat ttggttatga ttttggcacc agatgaattc caaaaaata tatatgaaac 96720
cgaaatcaag ccaaacttaa aaaccagtgc aattcttgca ttgacacatg gctttaatat 96780
tcatttttgaa aaaatagttc ctgaagcaac taacagcgtg attatgattg caccaaaaagg 96840
tccaggccat actgttagaa gtactttatac caatggtgga ggcgttccat ctctcatagc 96900
tatatatgaa gatgctttaa gtgatgaaga ttattcagca aaagatgtag ctctatctta 96960
tgcaaaagca aatggcggca cgagggctgg tgttcttgaa acatctttta aagaagaaac 97020
agaaacagat ttatttggcg aacaagcagt tttatgtggc gggcttaccg ctttaattaa 97080
agctggggtt gaaactctag ttgaggcagg ctacagtga gagatggcat attttgaatg 97140
ccttcatgaa acaaaactaa tcacagactt aattcaagaa ggtggcattg ctaatatgca 97200
ttactcaata tcaaatactg ctgagtatgg tgattatgtg agtggaccca aagtaattac 97260
tagcgatacc aaaaaagcaa tgaaaggaat actagaaaat atacagtcag gaaaattcgc 97320
agatgacttc ttgaatgact gtcgacaaaag caatgatggc actggcggac ctgttatgaa 97380
aagcaataga gaagctacaa aaattcatcc aatagagtgc gtaggggctg agctaaagtc 97440

taaatgaag ttctaaatt cacaaaaatt ggtggataaa gaaattaatt aaaattaatt 97500
aaaaaaaaagg tatcttcttc ggaataaactt cgttaagata cgcgtccgca caaagaggcg 97560
gttgttcttt aaaaatattt ggttactcgt gtgggtgttc aaatacagag aaaaaataat 97620
ttagattttt tataaaaatc acaaaaacat gatattaatt gaagagtttg atcatggctc 97680
agattgaacg ctggcggtag gcttaacaca tgcaagtcgt gcgagaaagt atcttcggat 97740
atgagtagag cggcggacgg gtgagtaacg cgtaggaatc tacctagtag aaggggatatag 97800
cccggggaaa cccggattaa taccgtatac ctcttcggg agaaagaagg cctctctttg 97860
aagctttcgc tactagatga gcctgcgtaa gattagcttg ttggtgaggt aaaggctcac 97920
caaggcgacg atctttagct ggtctgagag gacgatacgc cacattggga ctgagacacg 97980
gccagactc ctacgggagg cagcagtggg gaatattgga caatgggcgc aagcctgatac 98040
cagccatacc gcgtgtgtga agaaggcctt cgggttgtta agcactttaa gcaggagagaa 98100
aaagttataa gttaatacct tataaccctg atgttacctg cagaataagc accggctaata 98160
tccgtgccag cagccgcggt aatacggaag gtgcaagcgt taatcggaat tactgggcgt 98220
aaagcgcgcg tagtggtttt gttaagtgg atgtgaaagc cctgggctca acctaggaac 98280

tgcatccaaa actaactcac tagagtacga tagagggagg tagaattcat agtgtagcgg 98340
tggaatgcgt agatatattatg aagaatacca gtggcggaagg cggcctcctg gatctgtact 98400
gacactgagg tgcgaaagcg tgggtagcga acaggattag ataccctggt agtccacgcc 98460
gtaaacgatg acaactagct gtgggagac tatgtcttct agtgggcgag ctaacgcctt 98520
aagttgtccg cctggggagt acggccgcaa ggctaaaact caaatgaatt gacggggacc 98580
cgcacaagcg gtggagcatg tggtttaatt cgatgcaacg cgaaaaacct tacctactct 98640
tgacatactt ggaggctctt gtaatgagag tgtgcctttt ggaaccaaga tacagggtgct 98700
gcatggctgt cgtcagctcg tgtcgtgaga tgttccgtta agtcggataa cgagcgcaac 98760
ccttaccctt atttgccagc gattcggtcg ggaactataa ggggactgcc ggtgataaac 98820
cggagggaagg tgaggacgac gtcaagtcac catggccctt acgagtaggg ctacacacgt 98880
gctacaatgg gagatacaga cggacgctaa gccgcgaggt ggtgctaata ctaaaaagtc 98940
tttcgtagtc cggattggag tctgcaactc gactccatga agtcggaatc gctagtaatc 99000
gcggatcagc atgccgcggt gaatacgttc tcgggtcttg tacacacccg ccgtcacacc 99060
atgggaagtgg attgcaccag aagtagatag tctaacctta gggagggcgt ttaccacggt 99120

gtgcttcattg actggggtga agtcgtaaca aggtagccgt aggggaacct gtggctggat 99180
cacctcctta acgataaatc gcgtttttaa cgccacacag agtaatcaaa tattaaaaa 99240
aagaacattt agatatgtaa aatcatttgt atgtaatttt ctagtgtata catttatgta 99300
tacataagat cactgcaatt aaaaagtaac atatgcattt atgtgtatgt taaaaaagta 99360
attaatatat tttattaaagt tactctcaaa aatgaagata aaacttcaaa aaaaatatgt 99420
aacctttttt aaggttatat gatcaagtaa aggaagagca caaggcggat gccttggcag 99480
cataaggcga tgaaggacgt aataaccctgc gataagcctc ggggagctgg taaataagct 99540
tcgatccgag gatttccgaa tgggaaaacc caatacacat aagtgtatta tcttatactg 99600
aatacatagg tataagaggc aaacctaggg aactgaaaca tctaagtacc tagaggaaaa 99660
gaaatcaaca gagattccgg tagtagcggc gagcgaaacc ggaccagccc ttaagcttat 99720
tttagtccag caaatattc tggaaagtgt agccatagta ggtgatagcc ctgtatgcga 99780
aagactaatt taagtgaaat cgagtaggtc gggacacgag aaatctttgac tgaacatggg 99840
gggaccatcc tccaaggcta aatactctat gctgaccgat agtgaaccag taccgtgagg 99900
gaaaggcgaa aagaaccccg gcgaggggag tgaaatagaa cctgaaacct tgtgcttaca 99960

agcagtcgga gcagacttgt tctgtgacgg cgtacctttt gtataatggg tcaacgactt 100020
aattcagta gcaagcttaa ccatttaggg taggcgtagg gaaaccgagt cttaataggg 100080
cgctcagttg ctggaattag acccgaaacc gggtgatcta tccatggcca gtgtgaaggt 100140
cgagtaacat cgactggagg cgcgaaccca cttatgttga aaaatgaggg gatgagctgt 100200
ggataggagt gaaaggctaa tcaaacccgg agatagctgg ttctcttcga aaactattta 100260
ggtagtgctt cgtgtattac tgtagggggg agagcactgt ttcggcctagg gggtcacccc 100320
gacttacaa accgatgcaa actccgaata cctacaagta tgagcacggg agacagactg 100380
cgggtgctaa cgtccgtagt cgagagggaa acaaccaga ctgtcagcta aggtcccaaa 100440
ttatgattaa gtgggaaca atgtgggaag gcacaaacag ctaggagggtt ggcttagaag 100500
cagccatcct ttaaagaaag cgtaatagct cactagtcga gtcggcctgc gtggaagata 100560
taacgggggt aaatcataaa ccgaagctac agatcttaaa tttatttaag atggtagaag 100620
agcgttctgt aagcggttga aggtaagctg agaggcgaac tggacgtatc agaagtgcga 100680
atgttgacat gagtaacgat caaagaggtg aaaaacctct tcgccgaaaa accaagggtt 100740
cctgtccaac gctaatacga gcagggtgag gcggcccccta aggcgagggc gaaagccgta 100800

gtcgatggga aacagggttaa tattcctgta ctttttataa ctgcgatggg gtgacggaga 100860
aggttagact agcacggcga cggttgtccg tgttcaaggt tgtaggctgg tgttctaggt 100920
aatccggaa cgctaaggct gagaactgat aacgaccact ctacgagtgg gaagtagtgc 100980
ataccatgct tccaggaaaa acctctaagc ttcagggttat aagaaaccgt accctaaacc 101040
gacacagtg gttaggtcga gtagaccaag gtgtttgaga gaactatggt gaaggaaacta 101100
ggcaaaatag caccgtaact tcgggagaaag gtgcgccgcg tttgggtgat agacttgctc 101160
tctaagctga acgtggtcga agataccagg tggctgcgac tgtttactaa aaacatagca 101220
ctctgcaaac tcgtaagagg aagtataggg tgtgacgcct gcccggtgcc ggaagggttaa 101280
ttgatggggg tagcttatgc gaagctcttg atcgaagccc cggtaaacgg cggccgtaac 101340
tataacggtc ctaaggtagc gaaatcctt gtcgggtaag ttccgacctg cacgaatggc 101400
gtaacgatgg ccacactgtc tccaccatag actcagtga attgaaatcg ctgttaagat 101460
gcagtgtacc cgcagctaga cggaagacc ccgtgcacct ttactatagg ttcgcactgg 101520
actttgacct tacttggtga ggatagggtgg gagactttga agcagagacg ccagtctttg 101580
tggagtcac cttgaaatac cacccttgta agattgaagt tctaaccctag gtccattatc 101640

tggatcaggg acagtgcgtg ctgggtagtt tgactggggc ggtctctctc taaagagtaa 101700
cggaggagta cgaaggatc cttatcacgg tcggacatcg tgaggtaagt ataaaggcag 101760
aaggatgctt gactgcgaga tcgacggatc gagcaggtag gaaactaggt cttagtgatc 101820
cggtggttct gaatggaagg gccatcgctc aacggataaa aggtacgccg gggataaacag 101880
gctgataccg cccaagagtt catatcgacg gcgggtgtttg gcacctcgat gtcggctcat 101940
cacatcctgg ggctggagca ggtcccaagg gtatgggtgt tcgccattta aagtggtaag 102000
cgagctgggt ttagaacgtc gtgagacagt tcggtccta tctgctgtgg gcgtttggag 102060
atttgagga agctgattct agtacgagag gaccgaattg gacgaacctc tgggtgttccg 102120
gttgtcacgc cagtggcatt gccgggtagc tatgttcgga aaggataacc gctgaaagca 102180
tataagcggg aagcctctcc caagattaaa tctcccagag actttatgtc tcctaaagag 102240
tcgtcataga ctatgacgtt gataggcaag atgtgtaagc gctgcgaggc gttgagctaa 102300
cttgtactaa taactcgtga ggcttgatca tgtaacctta agcaaggttc ataatttgag 102360
taaaaacattg tagtgagaat taaaaataa aaagttacat accagtttgc ctgatgacaa 102420
tagcaacttg gaaccacctg atcccatctc gaactcagaa gtgaacgag ttaacgccaa 102480

tggtagtgca ggggtctccct gtgtgagagt aggaaatcgt caggcttttt tctttaaggc 102540
ttccagttta ctggaagcct ttttttttat ctcaagtata ataccagat gattattgga 102600
ttaacaggag gcatgggttc tggcaaatct gccgctgcag acttcttttat tgattttaggt 102660
atatcagtct tagatgcaga tcaagttgct aaagaagctt tatctacaaa ttctccttga 102720
tatactgatt ttatttctca atttggtgaa gtgtatttaa ataataatcg tgaggttgat 102780
aggctgaaat tgcgcgaact tattttttca aatccttcaa aaaaaaagg tcttgagaat 102840
attattcatc ccatagttcg gtctgctatt agtaatttta ttattacatc aacatcacca 102900
tattctattg ttatggtgcc actcattttt gaaacaaatt catataaaaa ttacgataag 102960
attattactg ttgactgtga tttagaactt cagatagtaa gggcctcaag tagagatgct 103020
caaaaataaat cgcagattaa gaatatattt aataagcaag cctctagaga ggagaggcta 103080
agtattttctg atgacgtact tatcaataac agcacccctat ctgatctaaa aaaacaagtt 103140
aatgttttac atactaaata tatggagtta ttaaatgagt agttgcccta gatgtgaaaa 103200
acctgtcaaa ctttctactg acaatattta tagaccttct tgctctgaga aatgcaaaact 103260
tatagatttt ggtgattggg ctaatgaaga taataaaatt tcaagaccacaa ttcaatctga 103320

agattttttac gaagattaaa tttaagaaag tctccattca ccacttttcaa tcatcggttt 103380
ggcttttttta tacttcatth cttgcgtatc ttgaccgtta gtaatthtta caagatcatt 103440
tctgcctaac ttaggttcgt ttcttgtcac agtacttggt ttactatag gagtggcttg 103500
tttttcattt tgaaagatat cagaattaat ttctctctttt tccaatthta attcttgaga 103560
actatttttgt ttgttttatgc tctctagctc agattctctgc gataatthgta atgagaaaaag 103620
tattcttata gtttcaacat ctattttccga aagcatagat tcaaacattg aatatgcctc 103680
tcttttgaat tcatttttttg gatttttttg agcataagcc cttagacca cactatttct 103740
taaatgatct atttctgata aatgctcttt ccaatggaca tcaagtactt gcagcataac 103800
ctgcttctca agaagcaatc tatttcacc aagatcacta aacttttttg aatatttatt 103860
ttttgcttgc agaacaattt cttcggcaat tgtatttggt acaagttttt tattactatt 103920
aattttattg gctatatctg ttcttagtcc atagctctct tttaaataat catcgagctg 103980
cctacttttc cattgagact ctattgattc ttccaggtag tataaattag atattccttt 104040
gaattgctgc tcgatgagt actcaatgggt actgctaata tcttctctct ctagtaattg 104100
attccttaaa gaatatatag cttgtctttg atcatttgat acatcgctgt actctaaaaag 104160

atctttcctt gcatcaaagt ttctgctctc aattcttttt tgcgcatttt ctattcctct 104220
tgaaagcatt ttagcttcaa tatgatcatc tcccatgccc agcctttcaa aagtagccct 104280
tcttccatcc gaaataaaaa gtctcaagag atcgctcttc aaagataaga agaattttga 104340
ataaacctgga tcaccttgtc tgcctgatct cccctttaac tgattatcta ttcttcttga 104400
ttcatgcctt tctgtttcaa gtatatgaag ccaccagat tctataactt ttttattatt 104460
ttctttccac tctatatcgc tttgggtcttc tttttttccg ccaagaacta tatccgtccc 104520
ccttcctgcc atattcgtag caatagttac cattccaggt ttacctgcat tggcaattat 104580
ctctgcttct ttttcgtgat gcttggcatt taaaatttga tgaggatttt tttgtttatt 104640
taaatatgct gatacttctt ctgaagatc aaccgaaact gttccaacaa gaatgggaga 104700
agattttttt cttaattggt caatttcttc aattagagct ttatatctcg attctgttgt 104760
aagaaaatact aagtcattaa gatcagctct agccatagga acatttgttg ggatgatgat 104820
gacatttagg ccatagattt gactaaactc tactgcttca gtatctgctg ttctgtctat 104880
cccagaaagt tttttaaata atctaaaaaa gttttggaat gtggtggatg ctagtgtttg 104940
agactctctt tggatagcaa catcttcttt gcattccagt gcctgggtgaa caccttcaact 105000

cattcttctt ccgggcattg ttctacctgt atgctcatca atcaaaagaa cctcacccgtt 105060
cctaaccaaa taatccacat tctttttaaa taagaagctt gctctaagtg ttgcttgaac 105120
aaatttcata atttttaaat tagaaacaga gtaagcccat ctgaggctcc aagccgattc 105180
cagc 105184

CLAIMS

What is claimed is:

1. A proteorhodopsin gene, comprising an isolated DNA sequence for encoding a proteorhodopsin protein.
2. The proteorhodopsin gene of claim 1, wherein said proteorhodopsin gene is retrieved from a genomic fragment of a sample of naturally occurring bacteria.
3. The proteorhodopsin gene of claim 2, wherein said naturally occurring bacteria are marine proteobacteria.
4. The proteorhodopsin gene of claim 2, wherein said naturally occurring bacteria are SAR86 bacteria.
5. The proteorhodopsin gene of claim 2, wherein said naturally occurring bacterial genomic fragment is retrieved from a recombinant DNA library.
6. The proteorhodopsin gene of claim 5, wherein said naturally occurring bacterial genomic fragment is retrieved from a bacterial artificial chromosome library.

7. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone BAC31A8, said proteorhodopsin gene is Sequence ID No:4 and said proteorhodopsin protein is Sequence ID No:5.
8. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone BAC40E8, said proteorhodopsin gene is Sequence ID No:8 and said proteorhodopsin protein is Sequence ID No:9.
9. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone BAC41B4, said proteorhodopsin gene is Sequence ID No:10 and said proteorhodopsin protein is Sequence ID No:11.
10. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone BAC64A5, said proteorhodopsin gene is Sequence ID No:12 and said proteorhodopsin protein is Sequence ID No:13.
11. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone HOT0m1, said proteorhodopsin gene is Sequence ID No:14 and said proteorhodopsin protein is Sequence ID No:15.

12. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone HOT75m1, said proteorhodopsin gene is Sequence ID No:16 and said proteorhodopsin protein is Sequence ID No:17.
13. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone HOT75m3, said proteorhodopsin gene is Sequence ID No:18 and said proteorhodopsin protein is Sequence ID No:19.
14. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone HOT75m4, said proteorhodopsin gene is Sequence ID No:20 and said proteorhodopsin protein is Sequence ID No:21.
15. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone HOT75m8, said proteorhodopsin gene is Sequence ID No:22 and said proteorhodopsin protein is Sequence ID No:23.
16. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB0m1, said proteorhodopsin gene is Sequence ID No:24 and said proteorhodopsin protein is Sequence ID No:25.

17. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB0m2, said proteorhodopsin gene is Sequence ID No:26 and said proteorhodopsin protein is Sequence ID No:27.
18. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB20m2, said proteorhodopsin gene is Sequence ID No:28 and said proteorhodopsin protein is Sequence ID No:29.
19. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB20m5, said proteorhodopsin gene is Sequence ID No:30 and said proteorhodopsin protein is Sequence ID No:31.
20. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB20m12, said proteorhodopsin gene is Sequence ID No:32 and said proteorhodopsin protein is Sequence ID No:33.
21. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB40m1, said proteorhodopsin gene is Sequence ID No:34 and said proteorhodopsin protein is Sequence ID No:35.

22. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB40m5, said proteorhodopsin gene is Sequence ID No:36 and said proteorhodopsin protein is Sequence ID No:37.
23. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB40m12, said proteorhodopsin gene is Sequence ID No:38 and said proteorhodopsin protein is Sequence ID No:39.
24. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB100m5, said proteorhodopsin gene is Sequence ID No:40 and said proteorhodopsin protein is Sequence ID No:41.
25. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB100m7, said proteorhodopsin gene is Sequence ID No:42 and said proteorhodopsin protein is Sequence ID No:43.
26. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB100m9, said proteorhodopsin gene is Sequence ID No:44 and said proteorhodopsin protein is Sequence ID No:45.

27. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone MB100m10, said proteorhodopsin gene is Sequence ID No:46 and said proteorhodopsin protein is Sequence ID No:47.
28. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALB1, said proteorhodopsin gene is Sequence ID No:48 and said proteorhodopsin protein is Sequence ID No:49.
29. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALB2, said proteorhodopsin gene is Sequence ID No:50 and said proteorhodopsin protein is Sequence ID No:51.
30. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALB5, said proteorhodopsin gene is Sequence ID No:52 and said proteorhodopsin protein is Sequence ID No:53.
31. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALB7, said proteorhodopsin gene is Sequence ID No:54 and said proteorhodopsin protein is Sequence ID No:55.

32. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALB6, said proteorhodopsin gene is Sequence ID No:56 and said proteorhodopsin protein is Sequence ID No:57.
33. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALB8, said proteorhodopsin gene is Sequence ID No:58 and said proteorhodopsin protein is Sequence ID No:59.
34. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALE1, said proteorhodopsin gene is Sequence ID No:60 and said proteorhodopsin protein is Sequence ID No:61.
35. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALE6, said proteorhodopsin gene is Sequence ID No:62 and said proteorhodopsin protein is Sequence ID No:63.
36. The proteorhodopsin gene of claim 2, wherein said genomic fragment is retrieved from a clone PALE7, said proteorhodopsin gene is Sequence ID No:64 and said proteorhodopsin protein is Sequence ID No:65.

37. The proteorhodopsin gene of claim 1, wherein said proteorhodopsin gene is amplified from a genomic fragment by polymerase chain reaction.
38. The proteorhodopsin gene of claim 37, wherein said polymerase chain reaction is performed by primers with Sequence ID No:2 and Sequence ID No:3.
39. The proteorhodopsin gene of claim 1, wherein said proteorhodopsin gene is derived from a marine environment and placed in an expression vector for producing said proteorhodopsin protein in a host.
40. The proteorhodopsin gene of claim 39, wherein said host is an artificial membrane system.
41. The proteorhodopsin gene of claim 39, wherein said host is a bacterium.
42. The proteorhodopsin gene of claim 41, wherein said host is a cell membrane preparation of said bacterium.
43. The proteorhodopsin gene of claim 39, wherein said host is an eukaryote.
44. The proteorhodopsin gene of claim 43, wherein said host is a cell membrane preparation of said eukaryote.

45. A method of retrieving a proteorhodopsin gene, comprising the steps of:
- (a) providing a sample of naturally occurring bacteria;
 - (b) extracting a genomic fragment of said sample of naturally occurring bacteria; and
 - (c) amplifying said proteorhodopsin gene from said genomic fragment using polymerase chain reaction.
46. The method of claim 45, further comprising the step of creating an expression vector containing said proteorhodopsin gene.
47. The method of claim 45, wherein said naturally occurring bacteria are marine proteobacteria.
48. The method of claim 45, wherein said naturally occurring bacteria are SAR86 bacteria.
49. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is retrieved from a recombinant DNA library.
50. The method of claim 49, said naturally occurring bacterial genomic fragment is retrieved from a bacterial artificial chromosome library.
51. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone BAC31A8, and wherein said amplified

proteorhodopsin gene from said clone BAC31A8 is Sequence ID No:4 and encodes a proteorhodopsin protein according to Sequence ID No:5.

52. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone BAC40E8, and wherein said amplified proteorhodopsin gene from said clone BAC40E8 is Sequence ID No:8 and encodes a proteorhodopsin protein according to Sequence ID No:9.
53. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone BAC41B4, and wherein said amplified proteorhodopsin gene from said clone BAC41B4 is Sequence ID No:10 and encodes a proteorhodopsin protein according to Sequence ID No:11.
54. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone BAC64A5, and wherein said amplified proteorhodopsin gene from said clone BAC64A5 is Sequence ID No:12 and encodes a proteorhodopsin protein according to Sequence ID No:13.
55. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone HOT0m1, and wherein said amplified proteorhodopsin gene from said clone HOT0m1 is Sequence ID No:14 and encodes a proteorhodopsin protein according to Sequence ID No:15.
56. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone HOT75m1, and wherein said amplified

proteorhodopsin gene from said clone HOT75m1 is Sequence ID No:16 and encodes a proteorhodopsin protein according to Sequence ID No:17.

57. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone HOT75m3, and wherein said amplified proteorhodopsin gene from said clone HOT75m3 is Sequence ID No:18 and encodes a proteorhodopsin protein according to Sequence ID No:19.
58. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone HOT75m4, and wherein said amplified proteorhodopsin gene from said clone HOT75m4 is Sequence ID No:20 and encodes a proteorhodopsin protein according to Sequence ID No:21.
59. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone HOT75m8, and wherein said amplified proteorhodopsin gene from said clone HOT75m8 is Sequence ID No:22 and encodes a proteorhodopsin protein according to Sequence ID No:23.
60. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB0m1, and wherein said amplified proteorhodopsin gene from said clone MB0m1 is Sequence ID No:24 and encodes a proteorhodopsin protein according to Sequence ID No:25.
61. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB0m2, and wherein said amplified

proteorhodopsin gene from said clone MB0m2 is Sequence ID No:26 and encodes a proteorhodopsin protein according to Sequence ID No:27.

62. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB20m2, and wherein said amplified proteorhodopsin gene from said clone MB20m2 is Sequence ID No:28 and encodes a proteorhodopsin protein according to Sequence ID No:29.
63. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB20m5, and wherein said amplified proteorhodopsin gene from said clone MB20m5 is Sequence ID No:30 and encodes a proteorhodopsin protein according to Sequence ID No:31.
64. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB20m12, and wherein said amplified proteorhodopsin gene from said clone MB20m12 is Sequence ID No:32 and encodes a proteorhodopsin protein according to Sequence ID No:33.
65. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB40m1, and wherein said amplified proteorhodopsin gene from said clone MB40m1 is Sequence ID No:34 and encodes a proteorhodopsin protein according to Sequence ID No:35.
66. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB40m5, and wherein said amplified

proteorhodopsin gene from said clone MB40m5 is Sequence ID No:36 and encodes a proteorhodopsin protein according to Sequence ID No:37.

67. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB40m12, and wherein said amplified proteorhodopsin gene from said clone MB40m12 is Sequence ID No:38 and encodes a proteorhodopsin protein according to Sequence ID No:39.
68. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB100m5, and wherein said amplified proteorhodopsin gene from said clone MB100m5 is Sequence ID No:40 and encodes a proteorhodopsin protein according to Sequence ID No:41.
69. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB100m7, and wherein said amplified proteorhodopsin gene from said clone MB100m7 is Sequence ID No:42 and encodes a proteorhodopsin protein according to Sequence ID No:43.
70. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB100m9, and wherein said amplified proteorhodopsin gene from said clone MB100m9 is Sequence ID No:44 and encodes a proteorhodopsin protein according to Sequence ID No:45.
71. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone MB100m10, and wherein said amplified

proteorhodopsin gene from said clone MB100m10 is Sequence ID No:46 and encodes a proteorhodopsin protein according to Sequence ID No:47.

72. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALB1, and wherein said amplified proteorhodopsin gene from said clone PALB1 is Sequence ID No:48 and encodes a proteorhodopsin protein according to Sequence ID No:49.
73. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALB2, and wherein said amplified proteorhodopsin gene from said clone PALB2 is Sequence ID No:50 and encodes a proteorhodopsin protein according to Sequence ID No:51.
74. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALB5, and wherein said amplified proteorhodopsin gene from said clone PALB5 is Sequence ID No:52 and encodes a proteorhodopsin protein according to Sequence ID No:53.
75. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALB7, and wherein said amplified proteorhodopsin gene from said clone PALB7 is Sequence ID No:54 and encodes a proteorhodopsin protein according to Sequence ID No:55.
76. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALB6, and wherein said amplified

proteorhodopsin gene from said clone PALB6 is Sequence ID No:56 and encodes a proteorhodopsin protein according to Sequence ID No:57.

77. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALB8, and wherein said amplified proteorhodopsin gene from said clone PALB8 is Sequence ID No:58 and encodes a proteorhodopsin protein according to Sequence ID No:59.

78. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALE1, and wherein said amplified proteorhodopsin gene from said clone PALE1 is Sequence ID No:60 and encodes a proteorhodopsin protein according to Sequence ID No:61.

79. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALE6, and wherein said amplified proteorhodopsin gene from said clone PALE6 is Sequence ID No:62 and encodes a proteorhodopsin protein according to Sequence ID No:63.

80. The method of claim 45, wherein said naturally occurring bacterial genomic fragment is in a clone PALE7, and wherein said amplified proteorhodopsin gene from said clone PALE7 is Sequence ID No:64 and encodes a proteorhodopsin protein according to Sequence ID No:65.

81. The method of claim 45, wherein said polymerase chain reaction is performed by primers with Sequence ID No:2 and Sequence ID No:3.

82. The method of claim 45, further comprising the step of providing a host.
83. The method of claim 82, wherein said host is an artificial membrane system.
84. The method of claim 82, wherein said host is a bacterium.
85. The method of claim 84, wherein said host is a cell membrane preparation of said bacterium.
86. The method of claim 82, wherein said host is an eukaryote.
87. The method of claim 86, wherein said host is a cell membrane preparation of said eukaryote.
88. A light-driven energy generator, comprising:
- (a) a proteorhodopsin protein;
 - (b) a host to correctly fold said proteorhodopsin protein in said host, thereby creating an integrated proteorhodopsin protein; and
 - (c) a source of retinal to bind covalently to said integrated proteorhodopsin protein, thereby creating a light absorbing pigment.
89. The light-driven energy generator of claim 88, wherein said proteorhodopsin protein is encoded by a proteorhodopsin gene retrieved from a genomic fragment of a sample of naturally occurring bacteria.

90. The light-driven energy generator of claim 89, wherein said naturally occurring bacteria are marine proteobacteria.
91. The light-driven energy generator of claim 89, wherein said naturally occurring bacteria are SAR86 bacteria.
92. The light-driven energy generator of claim 89, wherein said naturally occurring bacterial genomic fragment is retrieved from a recombinant DNA library.
93. The light-driven energy generator of claim 92, wherein said naturally occurring bacterial genomic fragment is retrieved from a bacterial artificial chromosome library.
94. The light-driven energy generator of claim 89, wherein said genomic fragment is retrieved from a clone, wherein said clone is a member of the group consisting of BAC31A8, BAC40E8, BAC41B4, BAC64A5, HOT0m1, HOT75m1, HOT75m3, HOT75m4, HOT75m8, MB0m1, MB0m2, MB20m2, MB20m5, MB20m12, MB40m1, MB40m5, MB40m12, MB100m5, MB100m7, MB100m9, MB100m10, PALB1, PALB2, PALB5, PALB7, PALB6, PALB8, PALE1, PALE6 and PALE7.
95. The light-driven energy generator of claim 88, wherein said host is an artificial membrane system.

96. The light-driven energy generator of claim 88, wherein said host is a cell membrane obtained from a bacterium.
97. The light-driven energy generator of claim 96, wherein said host is a cell membrane preparation obtained from a bacterium.
98. The light-driven energy generator of claim 88, wherein said host is a cell membrane obtained from an eukaryote.
99. The light-driven energy generator of claim 98, wherein said host is a cell membrane preparation obtained from an eukaryote.
100. The light-driven energy generator of claim 88, further comprising a light source for illuminating said light absorbing pigment, whereby said energy generator converts light into biochemical energy.
101. The light-driven energy generator of claim 100, wherein said light source is a fast-pulsed light source.
102. The light-driven energy generator of claim 101, wherein said fast-pulsed light source comprises a mechanism for delivering intermittent fast-light pulses at predetermined time intervals.

103. The light-driven energy generator of claim 100, wherein said light source is a light source exhibiting different predetermined wavelengths.
104. The light-driven energy generator of claim 88, further comprising a mediator for mediating energy generated by said energy generator into chemical, mechanical or electrical energy.
105. The light-driven energy generator of claim 88, wherein said proteorhodopsin protein is selected to determine an absorption spectra of said light absorbing pigment.
106. A method for making a light-driven energy generator, comprising the steps of:
- (a) providing a proteorhodopsin protein;
 - (b) providing a host to correctly fold said proteorhodopsin protein in said host, thereby creating an integrated proteorhodopsin protein; and
 - (c) providing a source of retinal to bind covalently to said integrated proteorhodopsin protein, thereby creating a light absorbing pigment.
107. The method of claim 106, wherein said proteorhodopsin protein is encoded by a proteorhodopsin gene retrieved from a genomic fragment of a sample of naturally occurring bacteria.
108. The method of claim 107, wherein said naturally occurring bacteria are marine proteobacteria.

109. The method of claim 107, wherein said naturally occurring bacteria are SAR86 bacteria.
110. The method of claim 107, wherein said naturally occurring bacterial genomic fragment is retrieved from a recombinant DNA library.
111. The method of claim 110, wherein said naturally occurring bacterial genomic fragment is retrieved from a bacterial artificial chromosome library.
112. The method of claim 107, wherein said genomic fragment is retrieved from a clone, wherein said clone is a member of the group consisting of BAC31A8, BAC40E8, BAC41B4, BAC64A5, HOT0m1, HOT75m1, HOT75m3, HOT75m4, HOT75m8, MB0m1, MB0m2, MB20m2, MB20m5, MB20m12, MB40m1, MB40m5, MB40m12, MB100m5, MB100m7, MB100m9, MB100m10, PALB1, PALB2, PALB5, PALB7, PALB6, PALB8, PALE1, PALE6 and PALE7.
113. The method of claim 106, wherein said host is an artificial membrane system.
114. The method of claim 106, wherein said host is a cell membrane obtained from a bacterium.

115. The method of claim 114, wherein said host is a cell membrane preparation obtained from a bacterium.
116. The method of claim 106, wherein said host is a cell membrane obtained from an eukaryote.
117. The method of claim 116, wherein said host is a cell membrane preparation obtained from an eukaryote.
118. The method of claim 106, further comprising the step of providing a light source for illuminating said light absorbing pigment, whereby said energy generator converts light into biochemical energy.
119. The method of claim 118, wherein said light source is a fast-pulsed light source.
120. The method of claim 119, wherein said fast-pulsed light source comprises a mechanism for delivering intermittent fast-light pulses at predetermined time intervals.
121. The method of claim 118, wherein said light source is a light source exhibiting different predetermined wavelengths.
122. The method of claim 106, further comprising the step of providing a mediator for mediating energy generated by said energy generator into chemical, mechanical or electrical energy.

123. The method of claim 106, wherein said proteorhodopsin protein is selected to determine an absorption spectra of said light absorbing pigment.
124. A PCR apparatus for amplifying a proteorhodopsin gene from DNA samples of naturally occurring microbial populations using polymerase chain reaction, comprising oligodeoxynucleotide primers with a Watson-Crick base pair complementarity to 5' and 3' ends of said proteorhodopsin gene.
125. The apparatus of claim 124, wherein said primers are according to Sequence ID No:2 and Sequence ID No:3.
126. A method of designing PCR primers, comprising the steps of:
- (a) determining a DNA sequence of a proteorhodopsin gene; and
 - (b) based on said determined DNA sequence in (a), designing oligodeoxynucleotide primers with a Watson-Crick base pair complementarity to said 5' and 3' ends of said proteorhodopsin gene.
127. The method of claim 126, further comprising the step of using said oligodeoxynucleotide primers to amplify said proteorhodopsin gene from DNA samples of naturally occurring microbial populations by polymerase chain reaction.
128. The method of claim 127, further comprising the step of cloning said amplified proteorhodopsin gene into an expression vector.

129. The method of claim 126, wherein said primers are according to Sequence ID No:2 and Sequence ID No:3.

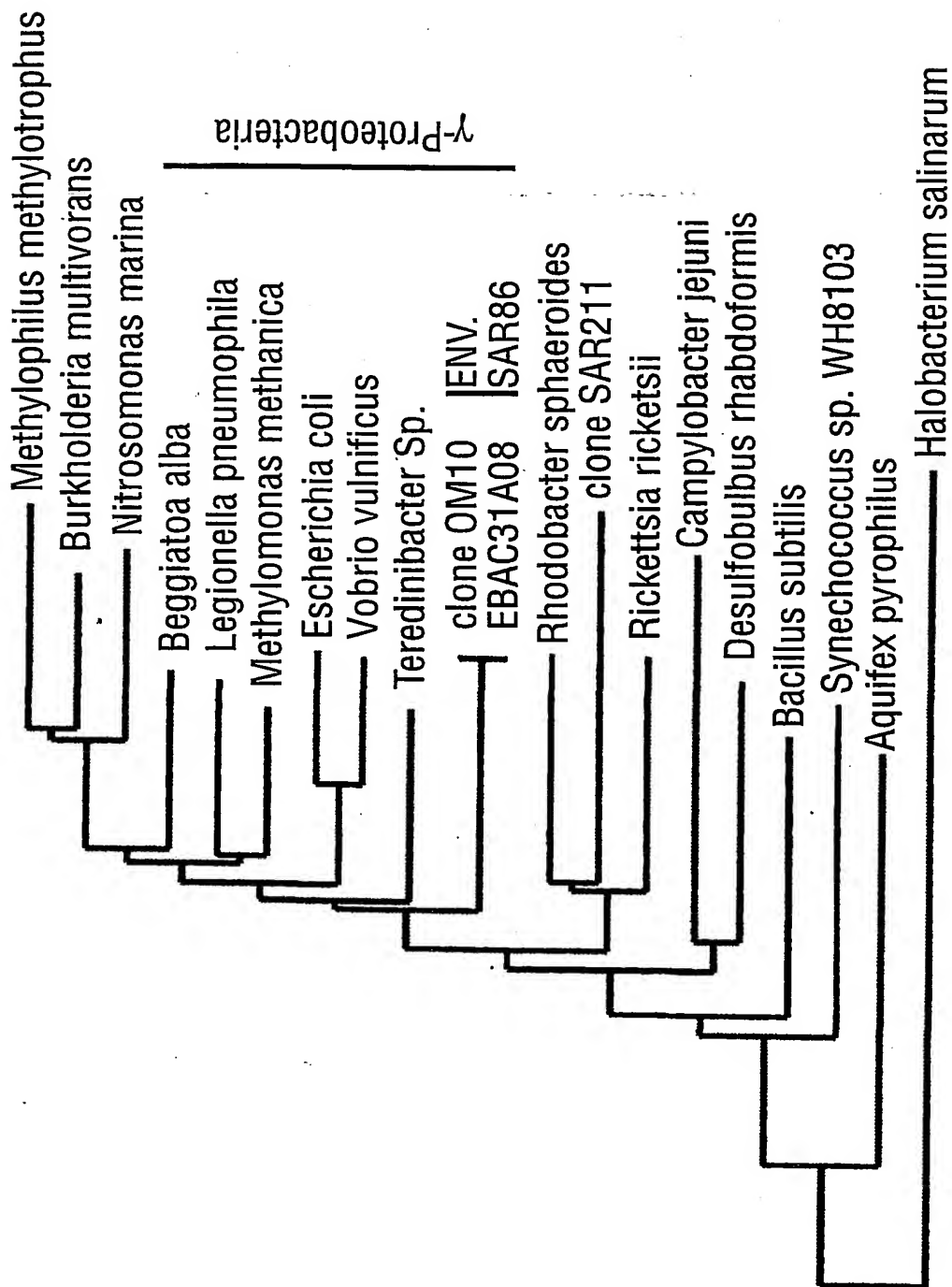


Fig. 1

29

accatgggta aattattact gatattagg

Figure 2

agcattagaa gattctttaa cagc 24

Figure 3

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca	48
Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr	
1 5 10 15	
ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt	96
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val	
20 25 30	
tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc	144
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe	
35 40 45	
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act	192
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr	
50 55 60	
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg	240
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met	
65 70 75 80	
aga ggg gta tgg att gaa act ggt gat tcg cca act act gta ttt aga tac	288
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr	
85 90 95	
att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta	336
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu	
100 105 110	

Figure 4

att ctt gct gct gca act aat gtt gct gga tca tta ttt aag aaa tta	384
Ile Leu Ala Ala gct gct gca act Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu	115 120 125
cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca	432
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala	130 135 140
gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg	480
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp	145 150 155 160
gta tac atg att tat gaa tta tgg gct gga gaa gga gaa tct gca tgt	528
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys	165 170 175
aat act gca agt cct gct gtg caa tca gct gct tac aac aca atg atg tat	576
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr	180 185 190
att atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt	624
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly	195 200 205
tac ctg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat	672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr	210 215 220

Figure 4

aac	ctt	gct	gac	ttt	gtt	aac	aag	att	cta	ttt	ggt	tta	att	ata	tgg	720
Asn	Leu	Ala	Asp	Phe	Val	Asn	Lys	Ile	Leu	Phe	Gly	Leu	Ile	Ile	Trp	
225					230					235					240	
aat	gtt	gct	gtt	aaa	gaa	tct	tct	aat	gct							750
Asn	Val	Ala	Val	Lys	Glu	Ser	Ser	Asn	Ala							
				245					250							

Figure 4

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca	48
Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr	15
1	
5	
10	
15	
20	
25	
30	
35	
40	
45	
50	
55	
60	
65	
70	
75	
80	
85	
90	
95	
100	
105	
110	
115	
120	
125	
130	
135	
140	
145	
150	
155	
160	
165	
170	
175	
180	
185	
190	
195	
200	
205	
210	
215	
220	
225	
230	
235	
240	
245	
250	
255	
260	
265	
270	
275	
280	
285	
290	
295	
300	
305	
310	
315	
320	
325	
330	
335	
340	
345	
350	
355	
360	
365	
370	
375	
380	
385	
390	
395	
400	
405	
410	
415	
420	
425	
430	
435	
440	
445	
450	
455	
460	
465	
470	
475	
480	
485	
490	
495	
500	
505	
510	
515	
520	
525	
530	
535	
540	
545	
550	
555	
560	
565	
570	
575	
580	
585	
590	
595	
600	
605	
610	
615	
620	
625	
630	
635	
640	
645	
650	
655	
660	
665	
670	
675	
680	
685	
690	
695	
700	
705	
710	
715	
720	
725	
730	
735	
740	
745	
750	
755	
760	
765	
770	
775	
780	
785	
790	
795	
800	
805	
810	
815	
820	
825	
830	
835	
840	
845	
850	
855	
860	
865	
870	
875	
880	
885	
890	
895	
900	
905	
910	
915	
920	
925	
930	
935	
940	
945	
950	
955	
960	
965	
970	
975	
980	
985	
990	
995	
1000	

Figure 6

att ctt gct gct gca aca aat gtt gct gct ggc ctg ttt aag aaa tta	384
Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu	
115 120 125	
ttg gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gag gca	432
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala	
130 135 140	
gga att atg aac gct tgg ggt gca ttc gtt att ggg tgt tta gct tgg	480
Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp	
145 150 155 160	
gta tac atg att tat gaa cta tgg gct gga gaa ggc aag gct gca tgt	528
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ala Ala Cys	
165 170 175	
aat act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat	576
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr	
180 185 190	
ata atc atc ttt ggt tgg gca att tat cct gta ggt tat ttc aca ggt	624
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly	
195 200 205	
tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat	672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Ile Tyr	
210 215 220	

Figure 6

gac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asp Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 6

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca	48
Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr	15
1	
5	
10	
20	
25	
30	
35	
40	
45	
50	
55	
60	
65	
70	
75	
80	
85	
90	
95	
100	
105	
110	
115	
120	
125	
130	
135	
140	
145	
150	
155	
160	
165	
170	
175	
180	
185	
190	
195	
200	
205	
210	
215	
220	
225	
230	
235	
240	
245	
250	
255	
260	
265	
270	
275	
280	
285	
290	
295	
300	
305	
310	
315	
320	
325	
330	
335	
340	
345	
350	
355	
360	
365	
370	
375	
380	
385	
390	
395	
400	
405	
410	
415	
420	
425	
430	
435	
440	
445	
450	
455	
460	
465	
470	
475	
480	
485	
490	
495	
500	
505	
510	
515	
520	
525	
530	
535	
540	
545	
550	
555	
560	
565	
570	
575	
580	
585	
590	
595	
600	
605	
610	
615	
620	
625	
630	
635	
640	
645	
650	
655	
660	
665	
670	
675	
680	
685	
690	
695	
700	
705	
710	
715	
720	
725	
730	
735	
740	
745	
750	
755	
760	
765	
770	
775	
780	
785	
790	
795	
800	
805	
810	
815	
820	
825	
830	
835	
840	
845	
850	
855	
860	
865	
870	
875	
880	
885	
890	
895	
900	
905	
910	
915	
920	
925	
930	
935	
940	
945	
950	
955	
960	
965	
970	
975	
980	
985	
990	
995	
1000	

Figure 7

aac ctt gct gat ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 7

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca	48
Met Gly Lys Leu Leu 5	
1	
ttt gct gca ggt ggc ggt gac gct ctt gat gct agt gat tac act ggt gtt	96
Phe Ala Ala Gly Gly 20	
25	
tct ttt tgg tta gtt aca gct gct gct cta tta gca tct act gta ttt ttc	144
Ser Phe Trp Leu Val 35	
40	
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act	192
Phe Val Glu Arg Asp Arg Val 55	
50	
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg	240
Val Ser Gly Leu Val 70	
75	
aga gga gta tgg att gaa act ggt gat tgg cct act gta ttt aga tac	288
Arg Gly Val Trp Ile 85	
90	
att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta	336
Ile Asp Trp Leu Leu 100	
105	
110	

Figure 8

att ctt gct gct gca act aat gtt gcc ggc tca tta ttt aag aaa ctt 384
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu
 115 120 125

cta gtt ggt tct ctt gtt gtt gtt gtt gtt ggt tac atg ggt gaa gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala
 130 135 140

gga att atg gca gct tgg cct gca ttc att att ggg tgt tta gct tgg 480
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp
 145 150 155 160

gta tac atg att tat gaa cta tat gct gga gaa gga gaa tct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys
 165 170 175

aat act gca agt cct tcg gtt caa tca gct tac aac aca atg atg gct 576
 Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala
 180 185 190

atc ata gtc ttc ggt ggt gca att tat cct ata ggt tat ttc aca ggt 624
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Ile Gly Tyr Phe Thr Gly
 195 200 205

tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt att tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Ile Tyr
 210 215 220

Figure 8

aac	ctt	gct	gac	ttt	gtt	aac	aag	att	cta	ttt	ggt	tta	att	ata	tgg	720
Asn	Leu	Ala	Asp	Phe	Val	Asn	Lys	Ile	Leu	Phe	Gly	Leu	Ile	Ile	Trp	
225					230					235					240	
aat	gtt	gct	gtt	aaa	gaa	tct	tct	aat	gct							750
Asn	Val	Ala	Val	Lys	Glu	Ser	Ser	Asn	Ala							
				245					250							

Figure 8

EBAC31A8	1	atgggtaaat	tattactgat	attaggTAGT	GTTATTGCAC	TTCCTACATT	50
EBAC40	1	50
EBAC41	1	50
EBAC64	1	50
EBAC31A8	51	TGCTGCAGGT	GGTGGTGACC	TTGATGCTAG	TGATTACACT	GGTGTTCCTT	100
EBAC40	51	100
EBAC41	51	100
EBAC64	51	C.....	100
EBAC31A8	101	TTTGGTTAGT	TACTGCTGCT	TTATTAGCAT	CTACTGTATT	TTTCTTTGTT	150
EBAC40	101	C.....	150
EBAC41	101	C.....	150
EBAC64	101	A.....	C.....	150
EBAC31A8	151	GAAAGAGATA	GAGTTTCTGC	AAAATGGAAA	ACATCATTA	CTGTATCTGG	200
EBAC40	151G..	200
EBAC41	151	200
EBAC64	151	200
EBAC31A8	201	TCTTGTTACT	GGTATTGCTT	TCTGGCATT	CATGTACATG	AGAGGGGTAT	250
EBAC40	201	250
EBAC41	201	250
EBAC64	201A....	250
EBAC31A8	251	GGATTGAAAC	TGGTGATTTCG	CCAACTGTAT	TTAGATACAT	TGATTGGTTA	300
EBAC40	251G..	300
EBAC41	251	300
EBAC64	251T.....	300

Figure 9

EBAC31A8	301	CTAACAGTTC	CTCTATTAAAT	ATGTGAATTC	TACTTAATTC	TTGCTGCTGC	350
EBAC40	301G.	350
EBAC41	301	350
EBAC64	301T.	350
EBAC31A8	351	AACTAATGTT	GCTGGATCAT	TATTTAAGAA	ATTACTAGTT	GGTTCCTTG	400
EBAC40	351	...A.....	...CTGGCC	...G.....	...T.G...	400
EBAC41	351	T.....	400
EBAC64	351C.C.C.T.	400
EBAC31A8	401	TTATGCTTGT	GTTTGGTTAC	ATGGGTGAAG	CAGGAATCAT	GGCTGCATGG	450
EBAC40	401G.T..	...AAC..T...	450
EBAC41	401	450
EBAC64	401T..	...A..T...	450
EBAC31A8	451	CCTGCATTCA	TTATTGGGTG	TTTAGCTTGG	GTATACATGA	TTTATGAATT	500
EBAC40	451	GG.....GC.	500
EBAC41	451C.	500
EBAC64	451C.	500
EBAC31A8	501	ATGGGCTGGA	GAAGGAAAAT	CTGCATGTAA	TACTGCAAGT	CCTGCTGTGC	550
EBAC40	501C.GG	550
EBAC41	501	550
EBAC64	501	..AT.....T.G..T.	550
EBAC31A8	551	AATCAGCTTA	CAACACAATG	ATGTATATTA	TCATCTTTGG	TTGGGCGGATT	600
EBAC40	551A.A...	600
EBAC41	551	600
EBAC64	551GC...C.	..AG....C..A...	600

Figure 9

EBAC31A8	601	TATCCTGTAG	GTTATTTCAC	AGGTTACCTG	ATGGGTGACG	GTGGATCAGC	650
EBAC40	601A	650
EBAC41	601	650
EBAC64	601A...A	650
EBAC31A8	651	TCCTTAACTTA	AACCTTATCT	ATAACCTTGC	TGACTTTGTT	AACAAGATTC	700
EBAC40	651G	700
EBAC41	651T	700
EBAC64	651T	700
EBAC31A8	701	TATTTGGTTT	AATTATATGG	AATGTTgctg	ttaaagaatc	ttctaagtct	750
EBAC40	701	750
EBAC41	701	750
EBAC64	701	750

Figure 9

EBAC31A8	1	MGKLLLIIGS	VIALPTFAAG	GGDLASDYT	GVSEWLVTA	LLASTVFFEV	50
EBAC40_1	1	50
EBAC41_1	1	A.....	50
EBAC64_1	1	50
EBAC31A8	51	ERDRVSAKWK	TSLTVSGLVT	GIAFWHYMM	RCVWIETGDS	PTVFRYIDWL	100
EBAC40_1	51	100
EBAC41_1	51	100
EBAC64_1	51	100
EBAC31A8	101	LTVPLLIICEF	YLILAAATNV	AGSLFKKLLV	GSLVMLVFGY	MGEAGIMAAW	150
EBAC40_1	101	AG.....N..	150
EBAC41_1	101	150
EBAC64_1	101	150
EBAC31A8	151	PAFIIGCLAW	VYMIYELWAG	EGKSACNTAS	PAVQSAYNMT	MYIIIFGWAI	200
EBAC40_1	151	G..V.....A.....	200
EBAC41_1	151	200
EBAC64_1	151Y..	S.....	A..V.....	200
EBAC31A8	201	YPVGYFTGYL	MGDGGGSA	NLIYNLADFV	NKILFGLIIW	NVAVKESNA	250
EBAC40_1	201D.....	250
EBAC41_1	201	250
EBAC64_1	201	I.....	250

Figure 10

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr
 1 5 10 15
 ttt gct gca ggt ggt ggt gac gct gct gct tta gca tct act gta ttt ttc 96
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
 20 25 30
 tct ttt tgg tta gtt act gct gct gct cta tta gca tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
 35 40 45
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tcg ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
 65 70 75 80
 aga ggg gta tgg att gag acc ggt gat tcg cca act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta cta aca gtt cct cta ttg ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
 100 105 110

Figure 11

att ctt gct gct gca aca aat gtt gct gct ggc ctg ttt aag aaa tta 384
 Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu
 115 120 125

ttg gtt ggt tct ctt gtt atg atg ggt ttt ggt tac atg ggt gag gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Phe Gly Tyr Met Gly Glu Ala
 130 135 140

gga att atg aac gct tgg ggt gca ttc gtt att ggg tgt tta gct tgg 480
 Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp
 145 150 155 160

gta tac atg att tat gaa cta tgg gct gga gaa ggc aag gct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ala Ala Cys
 165 170 175

aat act gca agt cct gct gct gct tca gct tac aac aca atg atg tat 576
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr
 180 185 190

ata atc atc ttt ggt tgg gca att tat cct gta ggt tat ttc aca ggt 624
 Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly
 195 200 205

tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr
 210 215 220

Figure 11

aac	ctt	gct	gac	ttt	gtt	aac	aag	att	cta	ttt	ggt	tta	att	ata	tgg	720
Asn	Leu	Ala	Asp	Phe	Val	Asn	Lys	Ile	Leu	Phe	Gly	Leu	Ile	Ile	Trp	
225					230					235					240	
aat	gtt	gct	gtt	aaa	gaa	tct	tct	aat	gct							750
Asn	Val	Ala	Val	Lys	Glu	Ser	Ser	Asn	Ala							
				245					250							

Figure 11

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
 1 5 10 15
 ttt gct gct gct ggt ggc gat gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
 20 25 30
 tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
 35 40 45
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt gct 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Ala
 50 55 60
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65 70 75 80
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
 100 105 110

Figure 12

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu
 115 120 125

 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala
 130 135 140

 gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp
 145 150 155 160

 tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val
 165 170 175

 agt act gca agt cct gct gct gtt aac tct gca tac aac gca atg atg atg 576
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met
 180 185 190

 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly
 195 200 205

 tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata 672
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile
 210 215 220

Figure 12

tat aac ctt gcc gac ctt gtt aac aag att cta ttt ggt ttg atc att
Tyr Asn Leu Ala Asp Leu Val Asn Lys Ile Leu Phe Gly Leu Ile
225 230 235 240 720

tgg aat gtt gct gtt gaa tct tct aat gct
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala
245 250 753

Figure 12

```

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca      48
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
1      5      10      15

ttt gct gct gct ggt ggc gat gta gat ata agt gat act gtt ggt gtt      96
Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
20      25      30

tca ttc tgg ctg ctg gtt aca gct gct ggt atg tta gcg gca act gta ttc ttt      144
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe
35      40      45

ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act      192
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50      55      60

gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tac atg      240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
65      70      75      80

aga ggt gtt tgg ata gat act ggt gat aca cca aca gta ttt aga tat      288
Arg Gly Val Trp Ile Asp Thr Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
85      90      95

att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta      336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
100      105      110

```

Figure 13

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu
 115 120 125 384

cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala
 130 135 140 432

ggt tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp
 145 150 155 160 480

tta tac atg att tat gag cta cat atg ggt gaa ggt aag gct gct gta
 Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Gly Lys Ala Ala Val
 165 170 175 528

agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg aag
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys
 180 185 190 576

att att gtt att gga tgg gca att tat cct gct gga tat gct gct ggt
 Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly
 195 200 205 624

tac cta atg agt ggt gac ggt gta tac gct gct tca aac tta aac ctt ata
 Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile
 210 215 220 672

Figure 13

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att	720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile	240
225	235
tgg aat gtt gct gtt aaa gaa gaa tct tct aat gct	753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala	250
245	

Figure 13

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca	48
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser	15
1	10
ttt gct gct gct ggt ggt ggc gat cta gat ata agt gat act gtt ggt gtt	96
Phe Ala Ala Ala Gly Gly Asp Ile Ser Asp Thr Val Gly Val	30
20	25
tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt	144
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe	45
35	40
ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act	192
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr	60
50	55
gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg	240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met	75
65	70
aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat	288
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr	90
85	95
att gat tgg tta act gtt cca tta caa gtg gtt gag ttc tat cta	336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Val Glu Phe Tyr Leu	110
100	105

Figure 14

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu
 115 120 125

 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala
 130 135 140

 gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480
 Gly Leu Ala Pro Val Leu Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp
 145 150 155 160

 tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Lys Ala Ala Val
 165 170 175

 agt act gca agt cct gct gct gtt aac tct gca tac aac gca atg atg atg 576
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met
 180 185 190

 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly
 195 200 205

 tac cta atg ggt ggc gaa ggt gta tac gct gct tca aac tta aac ctt ata 672
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile
 210 215 220

Figure 14

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att	720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile	240
225	
tgg aat gtt gct gtt gtt aaa gaa tct tct aat gct	753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala	250
245	

Figure 14

```

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca      48
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
1      5      10      15

ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt      96
Phe Ala Ala Ala Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
20      25      30

tca ttc tgg ctg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt      144
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
35      40      45

ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act      192
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50      55      60

gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg      240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
65      70      75      80

aga ggt gtt tgg ata gac act ggt gat acc cca gta ttc aga tat      288
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
85      90      95

att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta      336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
100      105      110

```

Figure 15

att ctt gct gct tgt aca aat gtt gct gct tca tta ttt aag aag ctt	384
Ile Leu Ala Ala Cys Thr Asn Val Ala Ala Ser Leu Phe Lys Lys Leu	
115 120 125	
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct	432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala	
130 135 140	
gga ttg gct cct gta tgg cct gct ttc att att ggt atg gct gga tgg	480
Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp	
145 150 155 160	
tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta	528
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Tyr Lys Ala Ala Val	
165 170 175	
agt act gca agt cct gct gct gtt aac tct gca tac aac gca atg atg gtg	576
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Val	
180 185 190	
att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt	624
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly	
195 200 205	
tac cta atg ggt ggc gaa ggt gta tac gct gct tca aac tta aac ctt ata	672
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	
210 215 220	

Figure 15

tat aac ctt gcc gac ctt gtt aac aag att cta ttt ggt ttg atc att 720
Tyr Asn Leu Ala Asp gac ctt gtt aac aag att cta ttt ggt ttg atc att
225 230 235 240
tgg aat gtt gct gtt gaa tct tct aat gct 753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala
245 250

Figure 15


```

48
atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca
Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr
1 5 10 15

96
ttt gct gca ggt ggt ggt gac ctt gat gct agt gat ggt ggt
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
20 25 30

144
tct ttt tgg tta gtt act gct gct gct cta tta gca tct act gta ttt ttc
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
35 40 45

192
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50 55 60

240
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
65 70 75 80

288
aga ggg gta tgg att gag act ggt gat tcg cca act gta ttt aga tac
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
85 90 95

336
att gat tgg tta cta aca gtt cct cta ttg ata tgt gaa ttc tac tta
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
100 105 110

```

Figure 16

```

att ctt gct gct gca aca aat gtt gct gct ggc ctg ttt aag aaa tta      384
Ile Leu Ala Ala 115
          120
          125

ttg gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gag gca      432
Leu Val Gly Ser Leu Val Met 135
          140
          145

gga att atg aac gct tgg cct gca ttc att att ggg tgt tta gct tgg      480
Gly Ile Met Asn Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp
145          150          155          160

gta tac atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt      528
Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys
165          170          175

aat act gca agt cct tcg gtt caa tca gct tac aac aca atg atg gct      576
Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala
180          185          190

atc ata gtc ttc ggt ggt gca att tat cct gta ggt tat ttc aca ggt      624
Ile Ile Val Phe 195
          200
          205

tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt att tat      672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr
210          215          220

```

Figure 16

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245 250

Figure 16

```

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca      48
Met Gly Lys Leu Leu 5 tta ggt agt ggt agt gtt att gca ctt cct aca
1                               10
                               15
ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt      96
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
20                               25
                               30
tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc      144
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
35                               40
                               45
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act      192
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50                               55
                               60
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg      240
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
65                               70
                               75
                               80
aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac      288
Arg Gly Val Trp 85 ggt Gly Asp Ser Pro Thr Val Phe Arg Tyr
                               90
                               95
att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta      336
Ile Asp Trp Leu Leu 100 Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
                               105
                               110

```

Figure 17

Figure 17

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 17

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr
 1 5 10 15
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat tac act ggt gtt 96
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
 20 25 30
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
 35 40 45
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
 65 70 75 80
 aga ggg gta tgg att gaa act ggt gat gct tgg cca act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
 100 105 110

Figure 18

att ctt gct gct gca act aat gtt gct gct ggc ctg ttt aag aaa tta 384
 Ile Leu Ala Ala Thr Asn Val Ala Ala Gly Leu Phe Lys Lys Leu 115 120 125

 ttg gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gag gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala 130 135 140

 gga att atg aac gct tgg ggt gca ttc gtt att ggg tgt tta gct tgg 480
 Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp 145 150 155 160

 gta tac atg att tat gaa cta tgg gct gga gaa ggc aag gct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ala Ala Cys 165 170 175

 aat act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat 576
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr 180 185 190

 ata atc atc ttt ggt tgg gca att tat cct gta ggt tat ttc aca ggt 624
 Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly 195 200 205

 tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr 210 215 220

Figure 18

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Trp 240
225 230 235

aat gtt gct gtt gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 18

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr
 1 5 10 15
 ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt 96
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
 20 25 30
 tct ttt tgg tta gtt aca gct gct gct cta tta gca tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu Ala Ser Thr Val Phe Phe
 35 40 45
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
 65 70 75 80
 aga ggg gta tgg att gaa act ggt gat tgg cca act act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
 100 105 110

Figure 19

Figure 19

aac ctt gct gac ttt gtt aac aag att cta ctt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Leu Gly Leu Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245 250

Figure 19

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu 5
 1
 ttt gct gca ggt ggt ggt gac ctt gat gct agt gat gtt ggt gtt 96
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
 20 25 30
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
 35 40 45
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
 65 70 75 80
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
 100 105 110

Figure 20

att ctt gct gct gca gct aat gtt gct gga tca tta ttt aag aaa tta 384
 Ile Leu Ala Ala Ala Val Ala Gly Ser Leu Phe Phe Lys Lys Leu
 115 120 125

 cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala
 130 135 140

 gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg 480
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp
 145 150 155 160

 gta tac atg att tat gaa tta tgg gct gga gaa gga aaa tct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys
 165 170 175

 aat act gca agt cct gct gtg caa tca gcc tac aac aca atg atg tat 576
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr
 180 185 190

 att atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt 624
 Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly
 195 200 205

 tac ttg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr
 210 215 220

Figure 20

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245 250

Figure 20

```

atg ggt aaa tta tta ctg ata ata ggt agt gtt att gca ctt cct aca      48
Met Gly Lys Leu Leu 5      15
1
ttt gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt      96
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
20 25 30
tct ttt tgg tta gtt tta gct gct gct cta tta gca tct act gta ttt ttc      144
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
35 40 45
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act      192
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50 55 60
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg      240
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
65 70 75 80
aga gga gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac      288
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
85 90 95
att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta      336
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
100 105 110

```

Figure 21

att ctt gct gct gca act aat gtt gcc ggc tca tta ttt aag aaa ctt 384
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu
 115 120 125
 cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala
 130 135 140
 gga att atg gca gct tgg cct gca ttc att att ggg tgt tta gct tgg 480
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp
 145 150 155 160
 gta tat atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys
 165 170 175
 aat aca gca agt cct gct gtg caa tca gct tac aac aca atg atg tat 576
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr
 180 185 190
 att atc gtc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt 624
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly
 195 200 205
 tac ctg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr
 210 215 220

Figure 21

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 21

```

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca      48
Met Gly Lys Leu Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr
1      5      10      15

ttt gct gca ggt ggt ggt gac-ctt gat gct agt gat tac act ggt ggt      96
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
20      25      30

tct ttt tgg tta gtt act gct gct cta tta gca tct act gta ttt ttc      144
Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
35      40      45

ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta act      192
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50      55      60

gta tcg ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg      240
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
65      70      75      80

aga ggg gta tgg att gag act ggt gat gct cca act gta ttt aga tac      288
Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
85      90      95

att gat tgg tta cta aca gtt cct cta ttg ata tgt gaa ttc tac tta      336
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
100      105      110

```

Figure 22

```

att ctt gct gct gca aca aat gtt gct gct ggc ctg ttt aag aaa tta      384
Ile Leu Ala Ala gct gct gct gct gct gct gct gct gct gct gct gct gct
115 120 125
ttg gtt ggt tct ctt gtt atg atg gtt ttt ggt tac atg ggt gag gca      432
Leu Val Gly Ser Leu Val Met Leu Val Phe Val Phe Gly Tyr Met Gly Glu Ala
130 135 140
gga att atg aac gct tgg ggt gca ttc gtt att ggg tgt tta gct tgg      480
Gly Ile Met Asn Ala Trp Gly Ala Phe Val Ile Gly Cys Leu Ala Trp
145 150 155
gta tac atg att tat gaa cta tgg gct gga gaa ggc aag gct gca tgt      528
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ala Ala Cys
165 170 175
aat act gca agt cct gct gct gtg caa tca gct tac aac aca atg atg tat      576
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr
180 185 190
ata atc atc ttt ggt tgg gca att tat cct gta ggt tat ttc aca ggt      624
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly
195 200 205
tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat      672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Asn Ile Tyr
210 215 220

```

Figure 22

aac ctt gct gac ttt gtt aac aag aat cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Asn Leu Phe Gly Leu Ile Ile Trp 240
225 230

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 22

```

atg ggt aaa tta tta tta cgg ata tta ggt agt gtt att gca ctt cct aca 48
Met Gly Lys Leu Leu 5 5 10 15
tct gct gca ggt ggc ggt gac ctt gat gct agt gat tac act ggt gtt 96
Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
20 25 30
tct ttt tgg tta gtt aca gct gct gct cta tta gca tct act gta ttt ttc 144
Ser Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
35 40 45
ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tta tta act 192
Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50 55 60
gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tat atg 240
Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
65 70 75 80
aga gga gta tgg att gaa act ggt gat tcg cca act act gta ttt aga tac 288
Arg Gly Val Trp Ile Glu Thr Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
85 90 95
att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta 336
Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
100 105 110

```

Figure 23

att ctt gct gct gca act aat gtt gct gga tca tta ttt aag aaa tta 384
 Ile Leu Ala Ala 115
 115
 cta gtt ggt tct ctt gtt atg ctt gtg ttt ggt tac atg ggt gaa gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala
 130 135 140
 gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg 480
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp
 145 150 155 160
 gta tac atg att tat gaa cta tgg gct gga gaa gga aaa tct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys
 165 170 175
 aat act gca agt cct gct gct gtg caa tca gct tac aac aca atg atg tat 576
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr
 180 185 190
 atc atc atc gtt ggt ggt gct ggt gct ggt ggt ggt tat ttc aca ggt 624
 Ile Ile Ile Val Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly
 195 200 205
 tac ctg atg ggt gac ggt gga tca gct ctt aac tta aac ctt atc tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr
 210 215 220

Figure 23

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 23

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu 5
 1
 ttt gct gca ggt ggc ggt gac .ctt gat gct agt gat gtt ggt gtt 96
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
 20 25 30
 tct ttt tgg tta gtt aca gct gct gct cta tta gca tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
 35 40 45
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt ctt gtt act ggt att gct gct ttc tgg cat tac atg tac atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
 65 70 75 80
 aga gga gta tgg att gaa act ggt gat tgg cca act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
 100 105 110

Figure 24

att ctt gct gct gca gct gtt gcc ggc tca tta ttt aag aaa ctt 384
 Ile Leu Ala Ala Thr Asn Val Ala Gly Ser Leu Phe Lys Lys Leu
 115 120 125

 cta gtt ggt tct ctt gtt atg ctt gtt ggt tac atg ggt gaa gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala
 130 135 140

 gga att atg gca gct tgg cct gca ttc att att ggg tgt tta gct tgg 480
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp
 145 150 155 160

 gta tac atg att tat gaa cta tat gct gga gaa gga aaa tct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Tyr Ala Gly Glu Gly Lys Ser Ala Cys
 165 170 175

 aat act gca agt cct tgg gtt caa tca gct tac aac aca atg atg gct 576
 Asn Thr Ala Ser Pro Ser Val Gln Ser Ala Tyr Asn Thr Met Met Ala
 180 185 190

 atc ata gtc ttc ggt ggt gca att tat cct gta ggt tat ttc aca ggt 624
 Ile Ile Val Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly
 195 200 205

 tac cta atg ggt gac ggt gga tca gct ctt aac tta aac ctt att tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr
 210 215 220

Figure 24

aac	ctt	gct	gac	ttt	ggt	aac	aag	att	cta	ttt	ggt	tta	att	ata	tgg	720
Asn	Leu	Ala	Asp	Phe	Val	Asn	Lys	Ile	Leu	Phe	Gly	Leu	Ile	Ile	Trp	
225					230					235					240	
aat	ggt	gct	ggt	aaa	gaa	tct	tct	aat	gct							750
Asn	Val	Ala	Val	Lys	Glu	Ser	Ser	Asn	Ala							
				245					250							

Figure 24

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu 5
 1
 ttt gct gca ggt ggt ggt gac gct gct gct tta gca tct act gta ttt ttc 96
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
 20 25 30
 tct ttt tgg tta gtt act gct gct gct tta gca tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
 35 40 45
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tac atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
 65 70 75 80
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta cta aca gtt cct cta tta ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
 100 105 110

Figure 25

[illegible]

Figure 25

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240
225 230 235

aat gct gct gtt aaa gaa tct.tct aat gct 750
Asn Ala Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 25

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu Ile Leu Gly Ser Val Ile Ala Leu Pro Thr
 1 5 10 15
 ttt gct gca ggt ggt ggt gac .ctt gat gct agt gat tac act ggt gtt 96
 Phe Ala Ala Gly Gly Gly Asp Leu Asp Ala Ser Asp Tyr Thr Gly Val
 20 25 30
 tct ttt tgg tta gtt act gct gct tta tta gca tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu Leu Ala Ser Thr Val Phe Phe
 35 40 45
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tat atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe Trp His Tyr Met Tyr Met
 65 70 75 80
 aga ggg gta tgg att gaa act ggt gat tcg cca act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Gly Asp Ser Pro Thr Val Phe Arg Tyr
 85 90 95
 ata gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Leu Ile Cys Glu Phe Tyr Leu
 100 105 110

Figure 26

att ctt gcc gct gca act aat gtt gct gga tca tta ttt aag aaa tta	384
Ile Leu Ala Ala gct gca act thr asn val ala gly ser leu phe lys leu	
115 120 125	
ctt gtt ggt tct ctt gtt atg .ctt gtg ttt ggt tac atg ggt gaa gca	432
Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala	
130 135 140	
gga atc atg gct gca tgg cct gca ttc att att ggg tgt tta gct tgg	480
Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Ile Gly Cys Leu Ala Trp	
145 150 155 160	
gta tac atg att tat gaa cta tgg gct gga gaa gga aaa tct gca tgt	528
Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Glu Lys Ser Ala Cys	
165 170 175	
aat act gca agt cct gct gtg caa tca gct tac aac aca atg atg tat	576
Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr	
180 185 190	
atc atc atc ttt ggt tgg gcg att tat cct gta ggt tat ttc aca ggt	624
Ile Ile Ile Phe Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly	
195 200 205	
tac ctt atg ggt gac ggt gga tca gca ctt aac tta aac ctt att tat	672
Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr	
210 215 220	

Figure 26

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Trp 240
225 230 235

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 26

atg ggt aaa tta tta ctg ata tta ggt agt gtt att gca ctt cct aca 48
 Met Gly Lys Leu Leu 5
 1
 ttt gct gca ggt ggc ggt ggt gac ctt gat gct agt gat tac act ggt gtt 96
 Phe Ala Ala Gly Gly Gly Asp Asp Leu 20
 20
 tct ttt tgg tta gtt aca gct gct gct cta tta gcg tct act gta ttt ttc 144
 Ser Phe Trp Leu Val Thr Ala Ala Leu 35
 35
 ttt gtt gaa aga gat aga gtt tct gca aaa tgg aaa aca tca tta act 192
 Phe Val Glu Arg Asp Arg Val Ser 50
 50
 gta tct ggt ctt gtt act ggt att gct ttc tgg cat tac atg tat atg 240
 Val Ser Gly Leu Val Thr Gly Ile Ala Phe 65
 65
 aga gga gta tgg att gaa act ggt gat tgg cca act gta ttt aga tac 288
 Arg Gly Val Trp Ile Glu Thr Gly Asp 85
 85
 att gat tgg tta cta aca gtt cct tta tta ata tgt gaa ttc tac tta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu 100
 100

Figure 27

att ctt gct gct gca act aat gtt gcc ggc tca tta ttt aag aaa ctt 384
 Ile Leu Ala Ala gct gca act thr Asn Val Ala gca Gly Ser Leu Phe Lys Lys Leu
 115 120 125
 cta gtt ggt tct ctt gtt atg .ctt gtg ttt ggt tac atg ggt gaa gca 432
 Leu Val Gly Ser Leu Val Met Leu Val Phe Gly Tyr Met Gly Glu Ala
 130 135 140
 gga ata atg gcg gct tgg cct gca ttc atc gtt gga tgt tta gca tgg 480
 Gly Ile Met Ala Ala Trp Pro Ala Phe Ile Val Gly Cys Leu Ala Trp
 145 150 155 160
 gta tat atg att tat gaa cta tgg gct ggt gaa gga aaa tct gca tgt 528
 Val Tyr Met Ile Tyr Glu Leu Trp Ala Gly Glu Gly Lys Ser Ala Cys
 165 170 175
 aat act gca agt cct gct gta cag tca gct tac aac aca atg atg tat 576
 Asn Thr Ala Ser Pro Ala Val Gln Ser Ala Tyr Asn Thr Met Met Tyr
 180 185 190
 atc atc atc gtt ggt tgg gca att tat cct gta ggt tat ttc aca ggt 624
 Ile Ile Ile Val Gly Trp Ala Ile Tyr Pro Val Gly Tyr Phe Thr Gly
 195 200 205
 tac cta atg ggt gac ggt gga tca gct ctt aat aac ctt att tat 672
 Tyr Leu Met Gly Asp Gly Gly Ser Ala Leu Asn Leu Ile Tyr
 210 215 220

Figure 27

aac ctt gct gac ttt gtt aac aag att cta ttt ggt tta att ata tgg 720
Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile Trp 240
225 230

aat gtt gct gtt aaa gaa tct tct aat gct 750
Asn Val Ala Val Lys Lys Glu Ser Ser Asn Ala 250
245

Figure 27

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu 5
 1
 ttt gct gct gct ggt ggc gat. cta gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala 20
 20
 tca ttc tgg ctg gtt aca gct gct ggt atg tta gcg gca act gta ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
 35
 40
 ttt gta gaa aga gac caa gtc agc gct agt aag tgg aaa act tca ctt act 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50
 55
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tac atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65
 70
 aga ggt gtt tgg ata gat act ggt gat gca cca aca gta ttt aga tat 288
 Arg Gly Val Trp Ile Asp Thr Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85
 90
 att gat tgg cta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
 100
 105
 110

Figure 28

att ctt gct gct tgt tga aca agt gtt gct gct tca tta ttt aag aag ctt 384
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu
 115 120 125

 cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala
 130 135 140

 ggt tta gct cct gta tta cct gct gct ttc att ctt ggt atg gct ggt tgg 480
 Gly Leu Ala Pro Val Leu Leu Pro Ala Phe Ile Leu Gly Met Ala Gly Trp
 145 150 155 160

 tta tac atg att tat gag cta cat atg ggt gaa ggt aag gct gct gta 528
 Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Gly Lys Ala Ala Val
 165 170 175

 agt act gca agt cct gct gtt aac tct gct tac aat gca atg atg aag 576
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys
 180 185 190

 att att gtt att gga tgg gca att tat cct gct gct tat gct gct ggt 624
 Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly
 195 200 205

 tac cta atg agt ggt gac ggt gta tac gct tca aac tta aac ctt ata 672
 Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Ile
 210 215 220

Figure 28

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att	720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile	240
225	230
tgg aat gtt gct gtt gaa gaa tct tct aat gct	753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala	250
245	250

Figure 28

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
 1 5 10 15
 ttt gct gct gct ggt ggc gat .cta gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
 20 25 30
 tca ttc tgg ctg gtt aca gct gct ggt atg tta gcg gca act gtg ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe
 35 40 45
 ttt gta gaa aga gac caa gtc agc gct gag tgg aaa act tca ctt act 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Glu Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65 70 75 80
 aga ggt gtt tgg ata gat act ggt gat acc cca aca gta ttc aga tat 288
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
 100 105 110

Figure 29

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Phe Leu Lys Lys Leu
 115 120

cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432
 Leu Ala Gly Ser Leu Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala
 130 135 140

gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480
 Gly Leu Ala Pro Val Leu Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp
 145 150 155 160

tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val
 165 170 175

agt act gca agt cct gct gct gtt aac tct gca tac aac gca atg atg atg 576
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met
 180 185 190

att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly
 195 200 205

tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata 672
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile
 210 215 220

Figure 29

tat aac ctt gct gac	ttt gtt aac aag att cta	ttt ggt ttg atc att	720
Tyr Asn Leu Ala Asp	Phe Val Asn Lys Ile Leu	Phe Gly Leu Ile	
225	230	240	
tgg aat gtt gct gtt	aaa gaa tct tct aat gct		753
Trp Asn Val Ala Val	Lys Glu Ser Ser Asn Ala		
	245	250	

Figure 29

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu 5
 1
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
 20 25 30
 tca ttc tgg ctg gtt aca gct gct ggt atg tta gcg gca act gtg ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
 35 40 45
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt tta att act ggt ata gcc ttt tgg cat tat ctg tat atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65 70 75 80
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
 100 105 110

Figure 30

att ctt gct gct tgt aca aat gtt gct gct gct tca tta ttt aag aag ctt	384
Ile Leu Ala Ala Cys Thr Asn Val Ala Ser Leu Phe Lys Lys Leu	
115 120 125	
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct	432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Glu Ala	
130 135 140	
gga tta gct cct gta tgg cct gct ttc att att ggt atg gct gga tgg	480
Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Gly Met Ala Gly Trp	
145 150 155 160	
tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta	528
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Lys Ala Ala Val	
165 170 175	
agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg	576
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met	
180 185 190	
att att gtt gtt gga tgg gca att tat cct gct gct gct gct ggt	624
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly	
195 200 205	
tac cta atg ggt ggc gaa ggt gta tac gct gct tca aac cta aac ctt ata	672
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	
210 215 220	

Figure 30

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att 720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile 240
225 230 235

tgg aat gtt gct gtt aaa gaa tct tct aat gct 753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 30

```

atg ggt aaa tta tta ctg ata tta ggt agt gct att gcg ctt cca tca      48
Met Gly Lys Leu Leu tta tta ggt agt gct att gcg ctt cca Pro Ser
1      5      10      15

ttt gct gct gct ggt ggc gat .cta gat ata agt gat act gtt ggt gtt      96
Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
20      25      30

tca ttc tgg ctg gtt acg gct gct ggt agt tta gcg gca act gta ttc ttt      144
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Ala Thr Val Phe Phe
35      40      45

ttt gta gaa aga gac caa gtc agc gct agt aag tgg aaa act tca ctt act      192
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
50      55      60

gta tct ggt tta att act ggt ata gct gct ttt tgg cat tat ctc tac atg      240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
65      70      75      80

aga ggt gtt tgg ata gat act ggt gat aca cca gta ttt aga tat      288
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
85      90      95

att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta      336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
100      105      110

```

Figure 31

att ctt gcc gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Phe Lys Lys Leu
 115 120 125
 cta gct ggt tca ttg gta atg tta ggt gct gga tct gca ggc gaa gct 432
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Ser Ala Gly Glu Ala
 130 135 140
 gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480
 Gly Leu Ala Pro Val Leu Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp
 145 150 155 160
 tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta 528
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Glu Lys Ala Ala Val
 165 170 175
 agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg 576
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met
 180 185 190
 att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly
 195 200 205
 tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctc ata 672
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile
 210 215 220

Figure 31

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att 720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Ile 240
225 230 235

tgg aat gtt gct gtt gaa tct tct aat gct 753
Trp Asn Val Ala Val Lys Glu Ser Asn Ala 250
245

Figure 31

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca	48
Met Gly Lys Leu Leu 5	
1	
ttt gct gct gct ggt ggt ggc gat gta gat act gtt ggt gtt	96
Phe Ala Ala 20	
25	
30	
tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt	144
Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe	
35	
40	
45	
ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act	192
Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr	
50	
55	
60	
gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctg atg	240
Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met	
65	
70	
75	
80	
aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat	288
Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr	
85	
90	
95	
att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta	336
Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu	
100	
105	
110	

Figure 32

att ctt gct gct tgt aca aat gtt gct gct tca tta ttt aag aag ctt	384
Ile Leu Ala Ala Cys Thr Asn Val Ala Ser Leu Phe Lys Lys Leu	
115 120 125	
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct	432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Glu Ala	
130 135 140	
gga tta gct cct gta tgg cct gct ttc att att ggt atg gct gga tgg	480
Gly Leu Ala Pro Val Trp Pro Ala Phe Ile Ile Glu Met Ala Gly Trp	
145 150 155 160	
tta tac atg att tat gag cta tat atg ggt gaa ggt aag gct gct gta	528
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Glu Lys Ala Ala Val	
165 170 175	
agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg gtg	576
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Val	
180 185 190	
att att gtt gga tgg gca att tat cct gct gga tat gct gct ggt	624
Ile Ile Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly	
195 200 205	
tac cta atg ggt ggc gaa ggt gta tac gct tca aac cta aac ctt ata	672
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	
210 215 220	

Figure 32

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att	720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile	240
225	235
tgg aat gtt gct gtt gtt aaa gaa tct tct aat gct	753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala	250
245	

Figure 32

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
 1 5 10 15
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala Gly Asp Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
 20 25 30
 tca ttc tgg ctg gtt aca gct gct ggt atg tta gcg gca act gtg ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
 35 40 45
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65 70 75 80
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
 100 105 110

Figure 33

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt	384
Ile Leu Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu	
115 120 125	
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct	432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala	
130 135 140	
gga tta gct cct gta tta cct gct ttc att ggt att gct gga tgg	480
Gly Leu Ala Pro Val Leu Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp	
145 150 155 160	
tta tac atg att tat gag cta tat atg ggt gaa ggt gct gct gta	528
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Tyr Lys Ala Ala Val	
165 170 175	
agt act gca agt gct gct gtt aac tct gca tac aac gca atg atg atg	576
Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met	
180 185 190	
att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt	624
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly	
195 200 205	
tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata	672
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	
210 215 220	

Figure 33

tat aac ctt gct gac ctt gtt aac aag att cta ttt ggt ttg atc att 720
Tyr Asn Leu Ala Asp Leu Val Asn Lys Ile Leu Phe Gly Leu Ile Ile 240
225 230 235

tgg aat gtt gct gtt gaa gaa tct tct aat gct 753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245 250

Figure 33

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu 5
 1
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala Ala 20
 30
 tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gtg ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
 35 40 45
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt tta att act act ggt ata gct ttt tgg cat tat ctc tat atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65 70 75 80
 aga ggt gtt tgg ata gac act ggt gat acc cca aca gta ttc aga tat 288
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta tta act gtt cca tta caa gtg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Val Val Glu Phe Tyr Leu
 100 105 110

Figure 34

att ctt gct gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt	384
Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu	
115 120 125	
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct	432
Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala	
130 135 140	
gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg	480
Gly Leu Ala Pro Val Leu Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp	
145 150 155 160	
tta tac atg att tat gag cta tat atg ggt gaa ggc aag gct gct gta	528
Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Gly Lys Ala Ala Val	
165 170 175	
agt act gca agt cct gct gtt aac cct gca tac aac gca atg atg atg	576
Ser Thr Ala Ser Pro Ala Val Asn Pro Ala Tyr Asn Ala Met Met Met	
180 185 190	
att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt	624
Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly	
195 200 205	
tac cta atg ggt ggc gaa ggt gta tac gct tca aac tta aac ctt ata	672
Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	
210 215 220	

Figure 34

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att 720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile 240
225 230 235

tgg aat gtt gct gtt gaa gaa gaa tct tct aat gct 753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 34

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu 5
 1 10 15
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala Gly Asp Gly Asp Leu Met 25
 20 30
 tca ttc tgg ctg gtt aca gct ggt atg tta gcg gca act gta ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
 35 40 45
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tac atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65 70 75 80
 aga ggt gtt tgg ata gat act ggt gat aca cca aca gta ttt aga tat 288
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
 100 105 110

Figure 35

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu	115 120 125	384
cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Gly Glu Ala	130 135 140	432
ggt tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp	145 150 155 160	480
tta tac atg att tat gag cta cat atg ggt gaa ggt aag gct gct gta Leu Tyr Met Ile Tyr Glu Leu His Met Gly Glu Gly Lys Ala Ala Val	165 170 175	528
agt act gca agt cct gct gct gtt aac tct gca tac aac gca atg atg aag Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Lys	180 185 190	576
att att gtt att gga tgg gca att tat cct gct gga tat gct gct ggt Ile Ile Val Ile Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly	195 200 205	624
tac cta atg agt ggt gac ggt gta tac gct tca aac tta aac ctt ata Tyr Leu Met Ser Gly Asp Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile	210 215 220	672

Figure 35

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att 720
Tyr Asn Leu Ala Asp Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile 240
225 230 235

tgg aat gtt gct gtt aaa gaa tct tct aat gct 753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala 250
245

Figure 35

atg ggt aaa tta tta ctg ata tta ggt agt gct att gca ctt cca tca 48
 Met Gly Lys Leu Leu Ile Leu Gly Ser Ala Ile Ala Leu Pro Ser
 1 5 10 15
 ttt gct gct gct ggt ggc gat cta gat ata agt gat act gtt ggt gtt 96
 Phe Ala Ala Ala Gly Gly Asp Leu Asp Ile Ser Asp Thr Val Gly Val
 20 25 30
 tca ttc tgg ctg ctg gtt aca gct gct ggt atg tta gcg gca act gtg ttc ttt 144
 Ser Phe Trp Leu Val Thr Ala Gly Met Leu Ala Thr Val Phe Phe
 35 40 45
 ttt gta gaa aga gac caa gtc agc gct aag tgg aaa act tca ctt act 192
 Phe Val Glu Arg Asp Gln Val Ser Ala Lys Trp Lys Thr Ser Leu Thr
 50 55 60
 gta tct ggt tta att act ggt ata gct ttt tgg cat tat ctc tat atg 240
 Val Ser Gly Leu Ile Thr Gly Ile Ala Phe Trp His Tyr Leu Tyr Met
 65 70 75 80
 aga ggt gtt tgg ata gat act ggt gat acc cca aca gta ttc aga tat 288
 Arg Gly Val Trp Ile Asp Thr Gly Asp Thr Pro Thr Val Phe Arg Tyr
 85 90 95
 att gat tgg tta tta act gtt cca tta caa atg gtt gag ttc tat cta 336
 Ile Asp Trp Leu Leu Thr Val Pro Leu Gln Met Val Glu Phe Tyr Leu
 100 105 110

Figure 36

att ctt gct gct tgt aca agt gtt gct gct tca tta ttt aag aag ctt 384
 Ile Leu Ala Ala Cys Thr Ser Val Ala Ala Ser Leu Phe Lys Lys Leu
 115 120 125

cta gct ggt tca tta gta atg tta ggt gct gga ttt gca ggc gaa gct 432
 Leu Ala Gly Ser Leu Val Met Leu Gly Ala Gly Phe Ala Glu Ala
 130 135 140

gga tta gct cct gta tta cct gct ttc att att ggt atg gct gga tgg 480
 Gly Leu Ala Pro Val Leu Pro Ala Phe Ile Ile Gly Met Ala Gly Trp
 145 150 155 160

cta tac atg att tat gag cta tat atg ggt gaa ggt gct gct gta 528
 Leu Tyr Met Ile Tyr Glu Leu Tyr Met Gly Glu Glu Lys Ala Ala Val
 165 170 175

agt act gca agt cct gct gtt aac tct gca tac aac gca atg atg atg 576
 Ser Thr Ala Ser Pro Ala Val Asn Ser Ala Tyr Asn Ala Met Met Met
 180 185 190

att att gtt gtt gga tgg gca att tat cct gct gga tat gct gct ggt 624
 Ile Ile Val Val Gly Trp Ala Ile Tyr Pro Ala Gly Tyr Ala Ala Gly
 195 200 205

tac cta atg ggt ggc gaa ggc gta tac gct gct tca aac tta aac ctt ata 672
 Tyr Leu Met Gly Gly Glu Gly Val Tyr Ala Ser Asn Leu Asn Leu Ile
 210 215 220

Figure 36

tat aac ctt gct gac ttt gtt aac aag att cta ttt ggt ttg atc att	720
Tyr Asn Leu Ala Asp Phe Phe Val Asn Lys Ile Leu Phe Gly Leu Ile Ile	240
225	230
tgg aat gtt gct gtt gaa gaa tct tct aat gct	753
Trp Asn Val Ala Val Lys Glu Ser Ser Asn Ala	250
245	

Figure 36

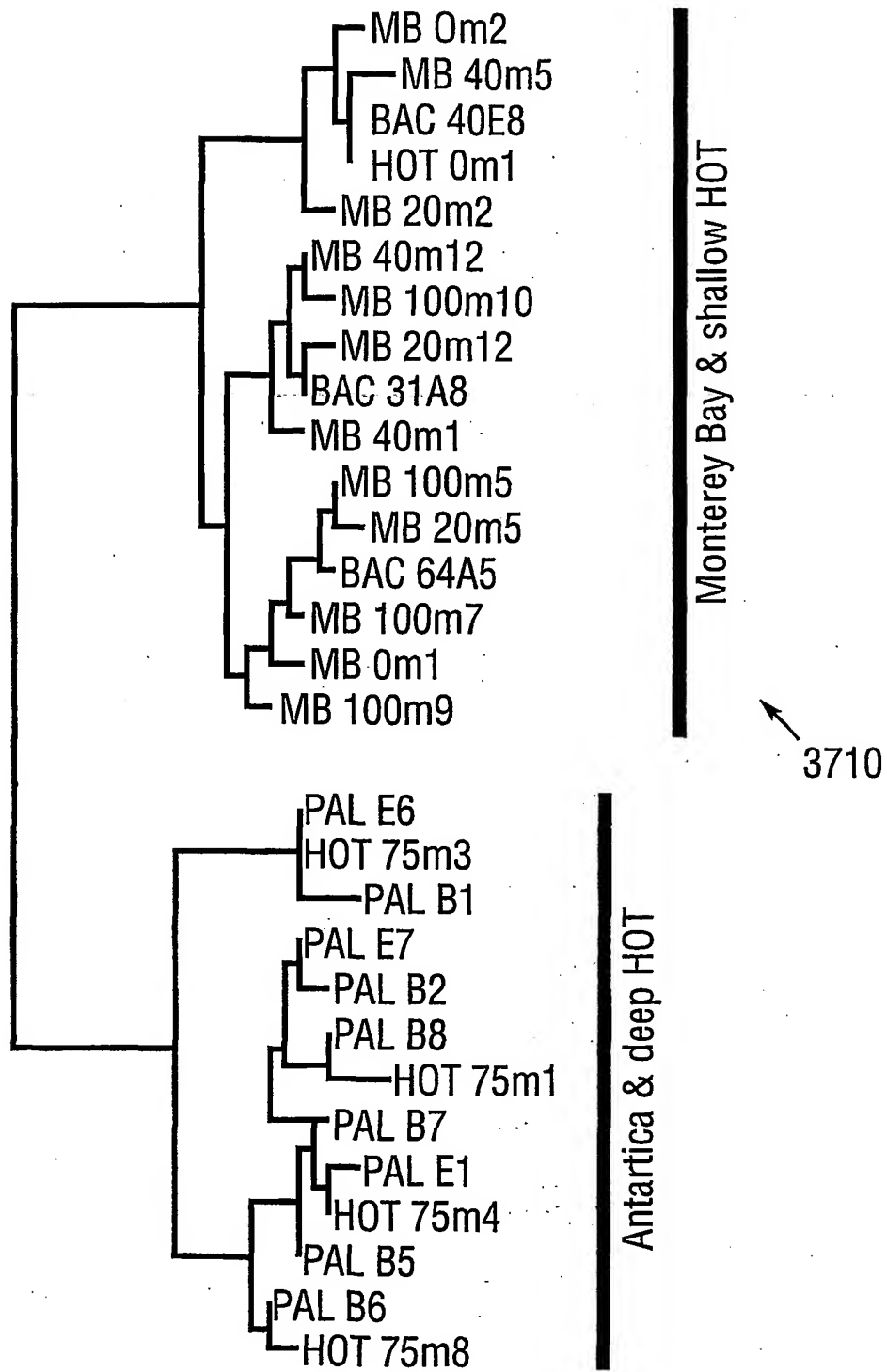
**Fig. 37**



Fig. 38A

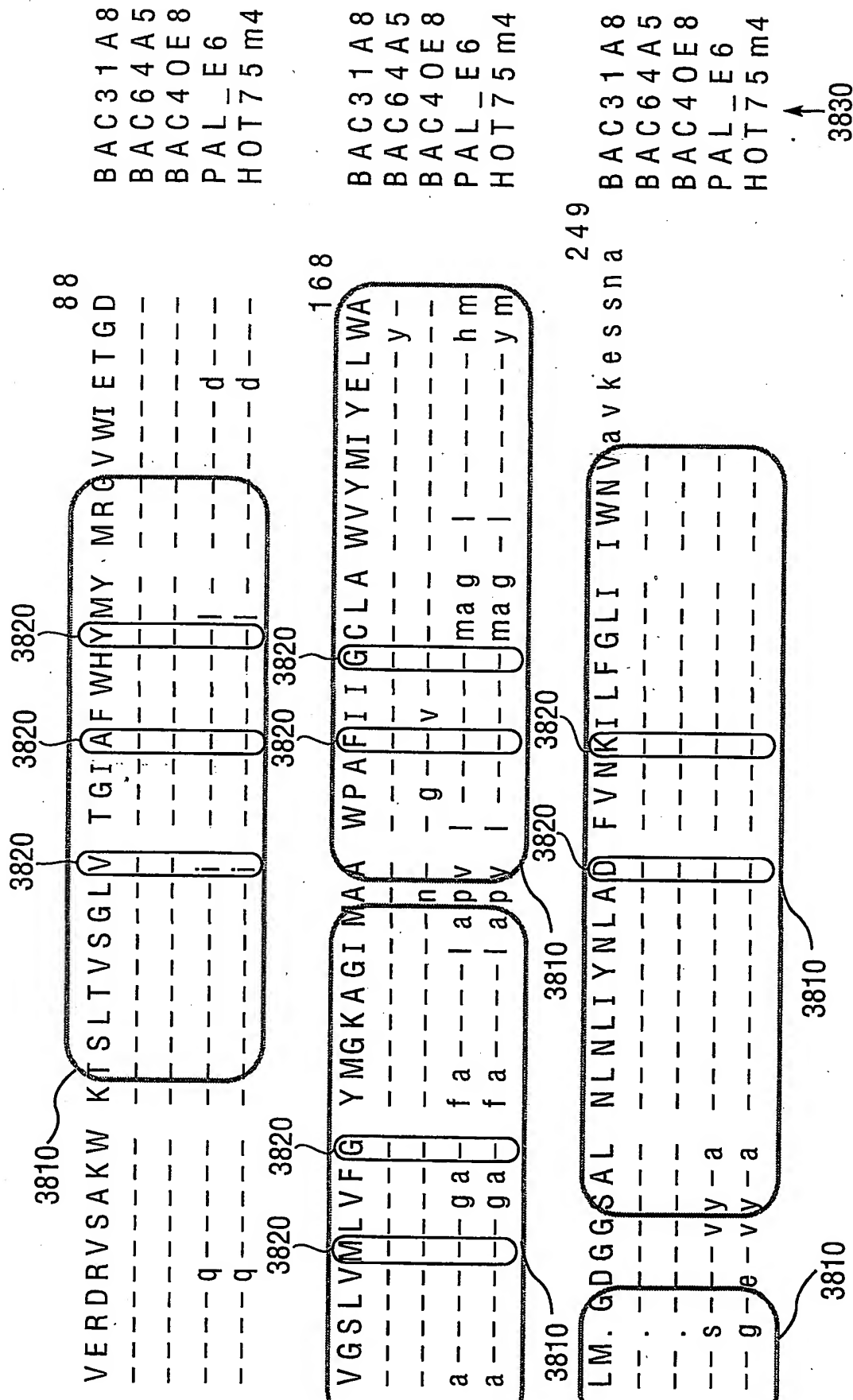


Fig. 38B

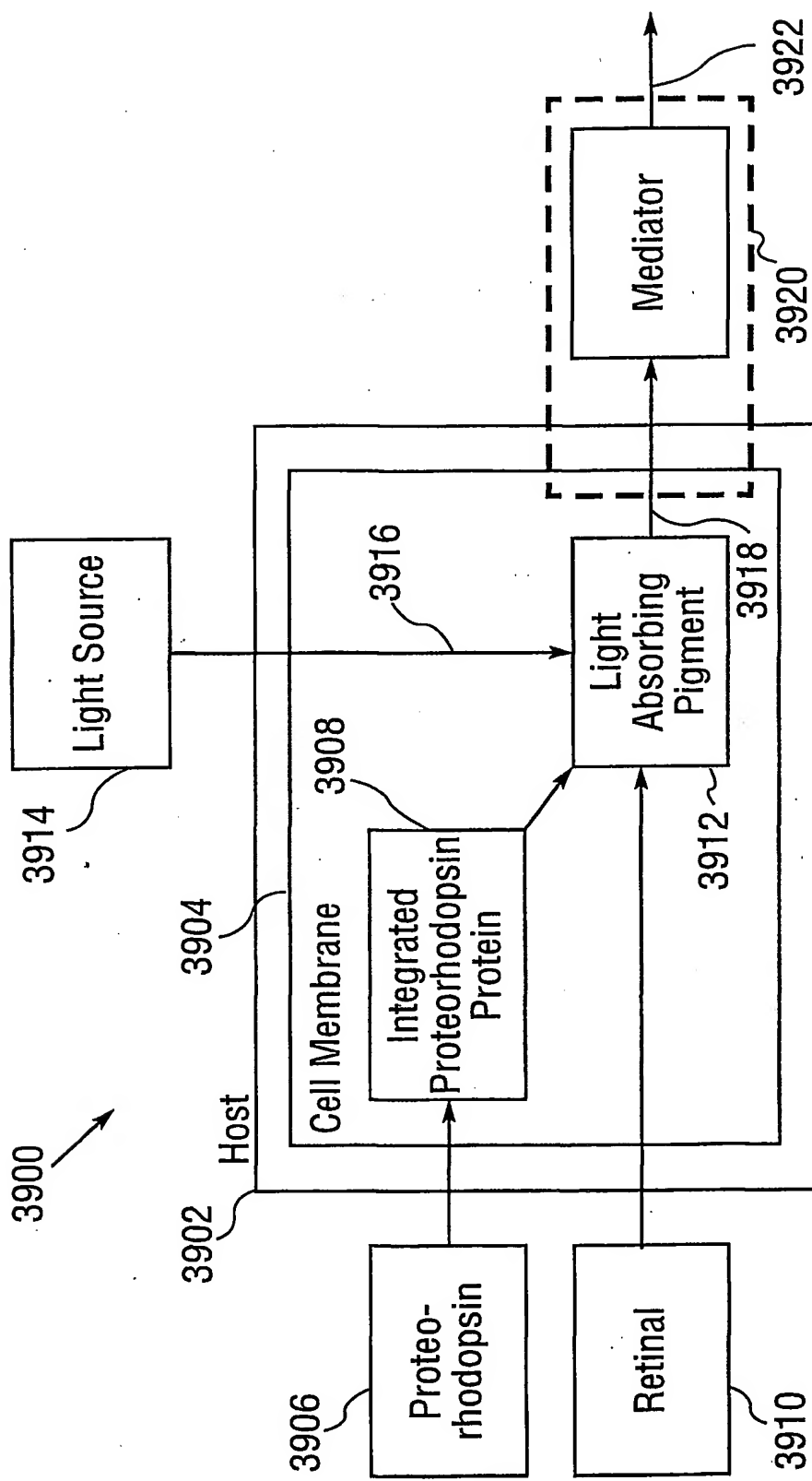
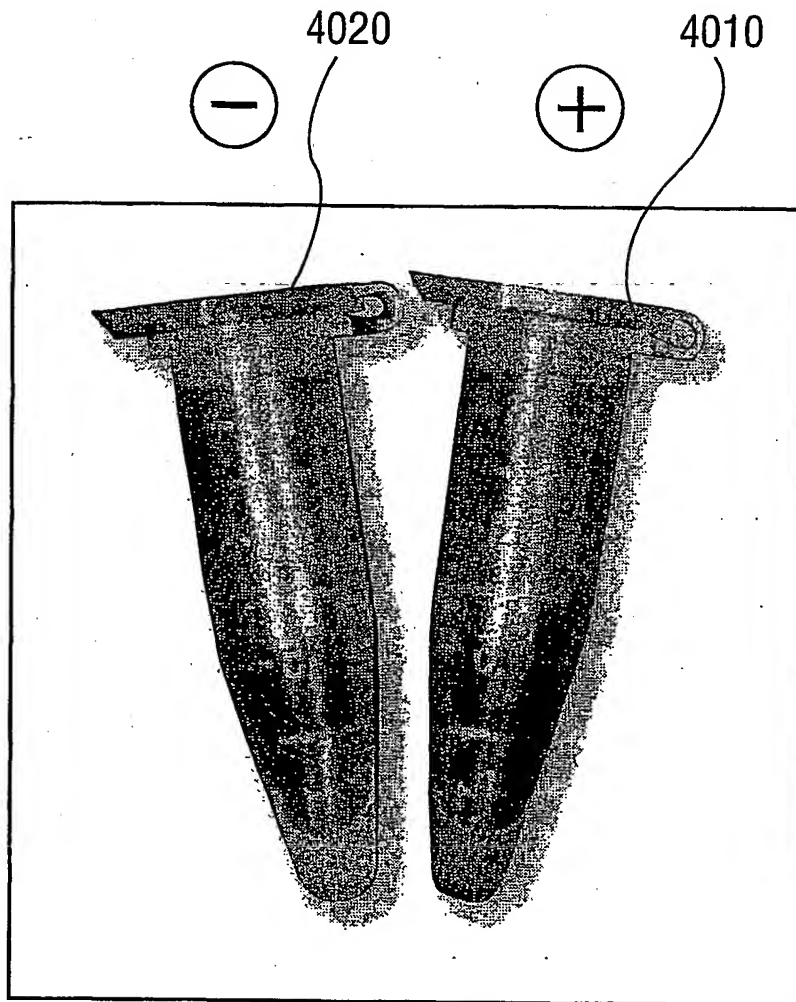
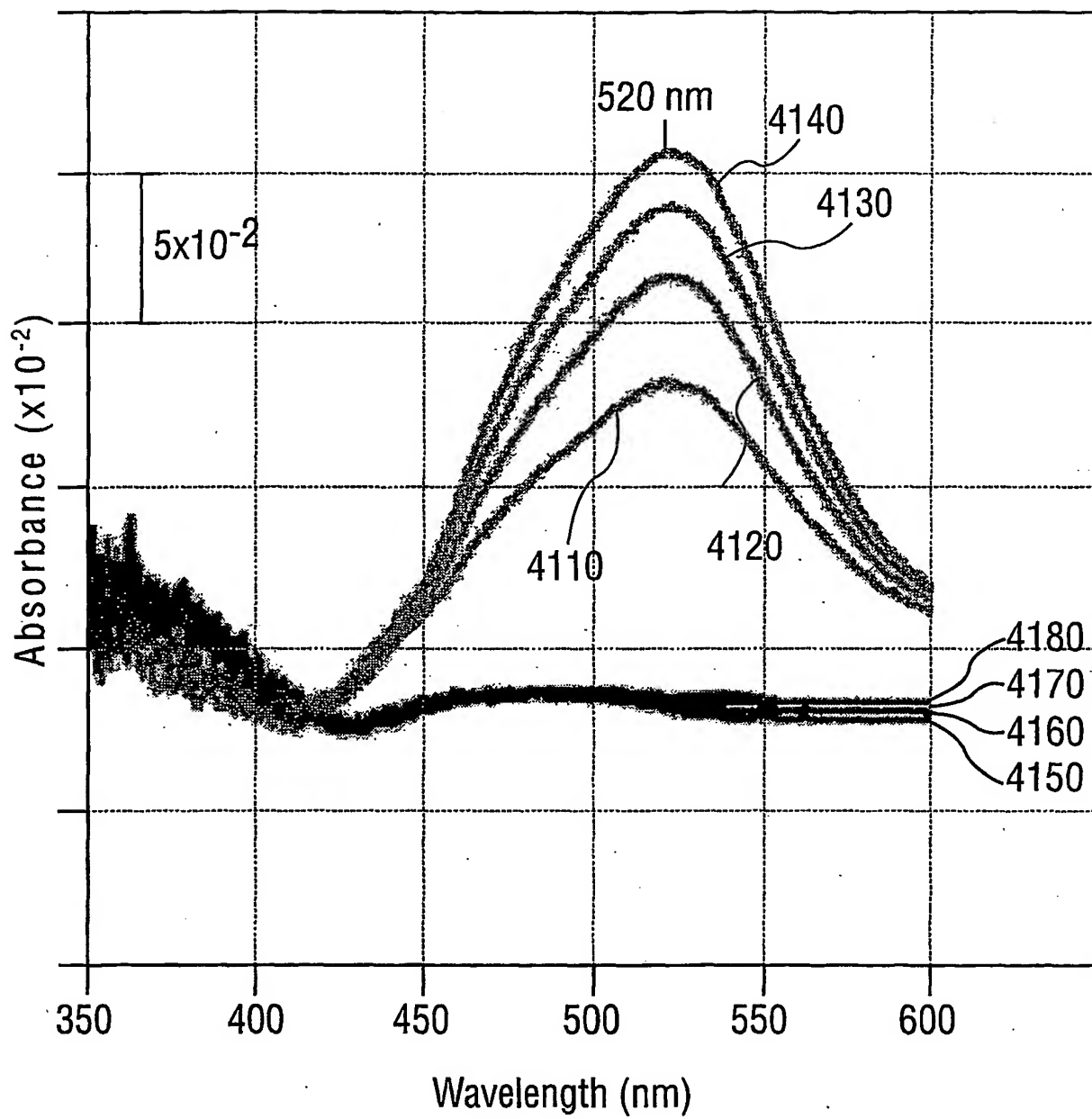


Fig. 39

**Fig. 40**

**Fig. 41**

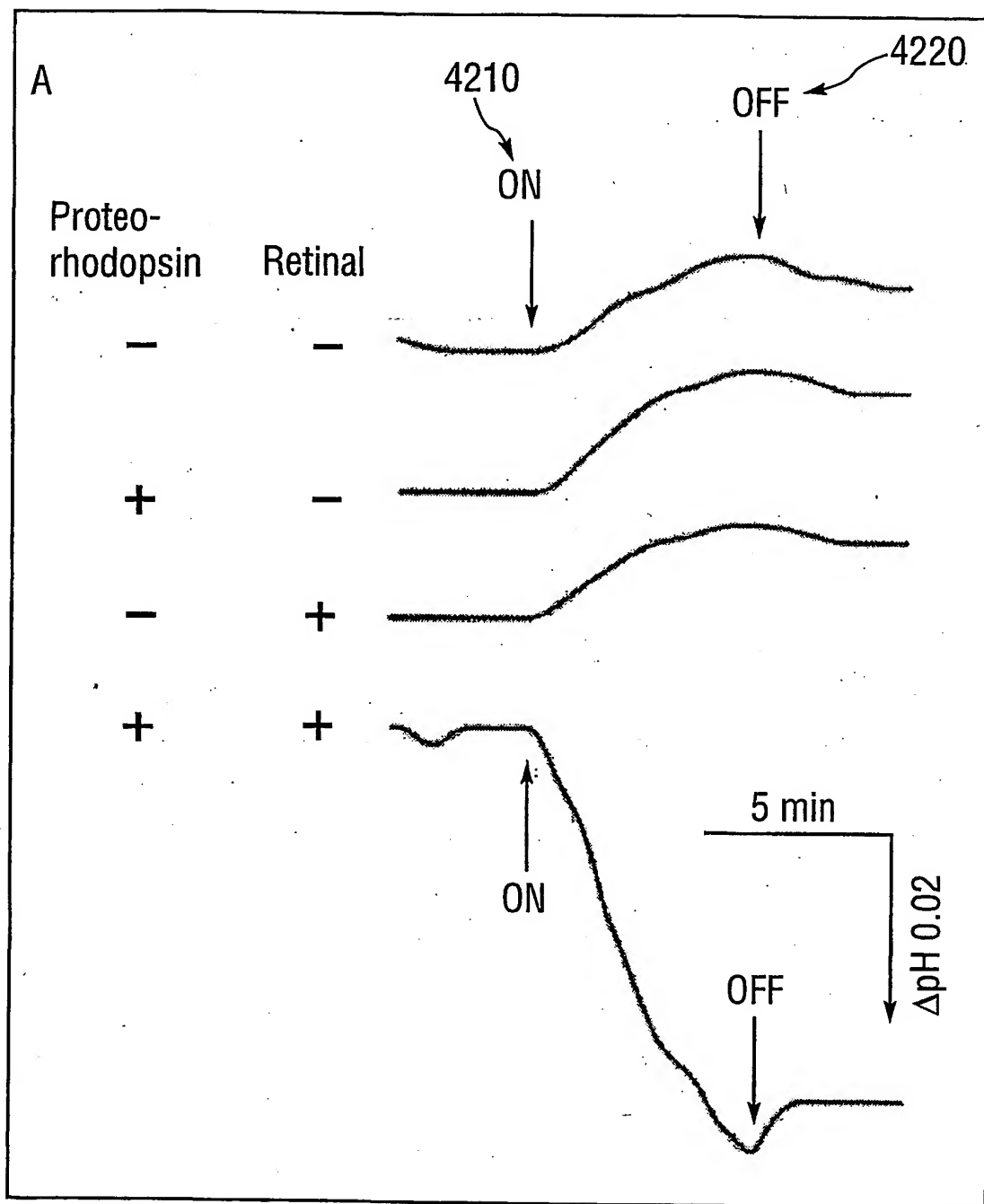


Fig. 42

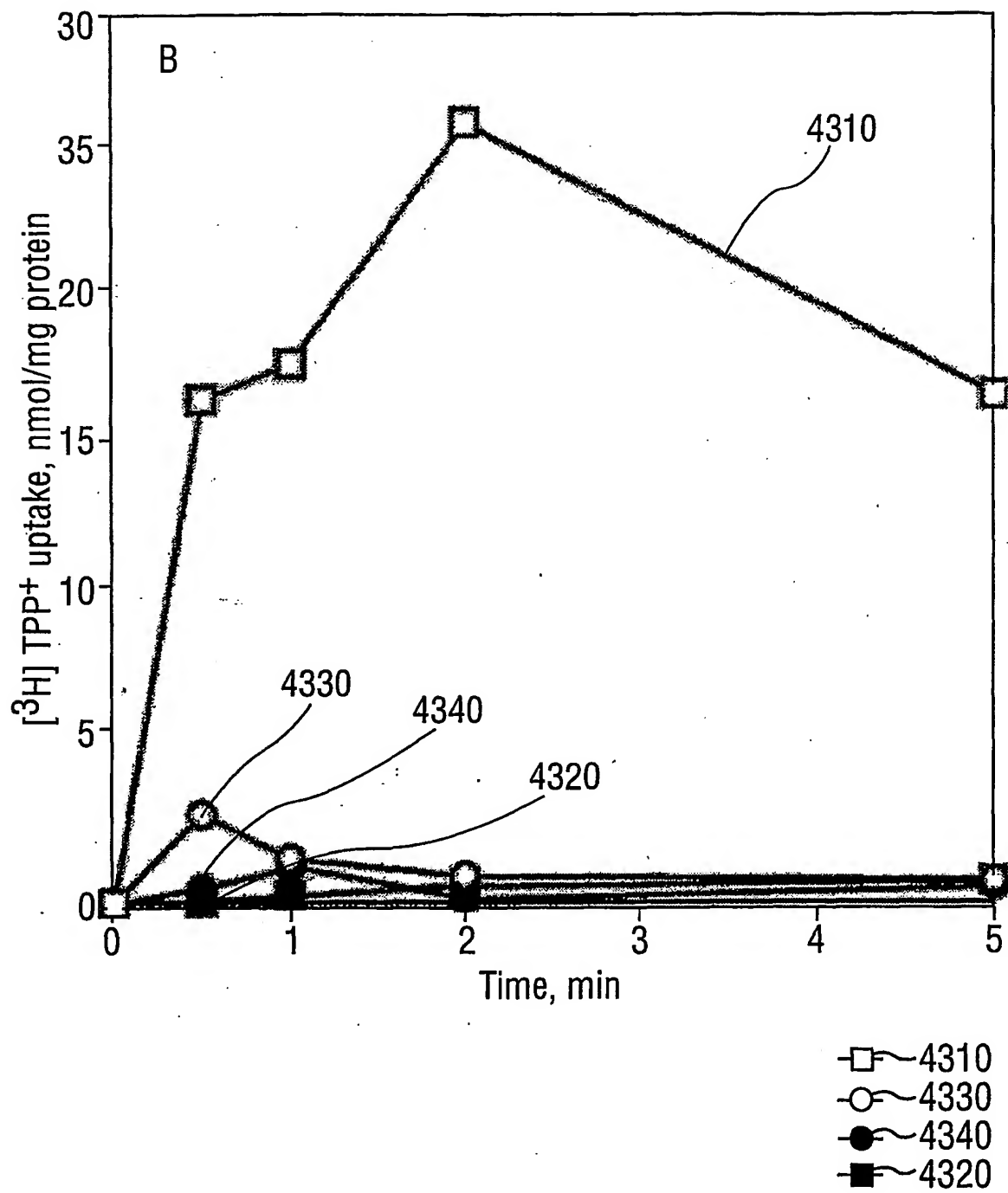


Fig. 43

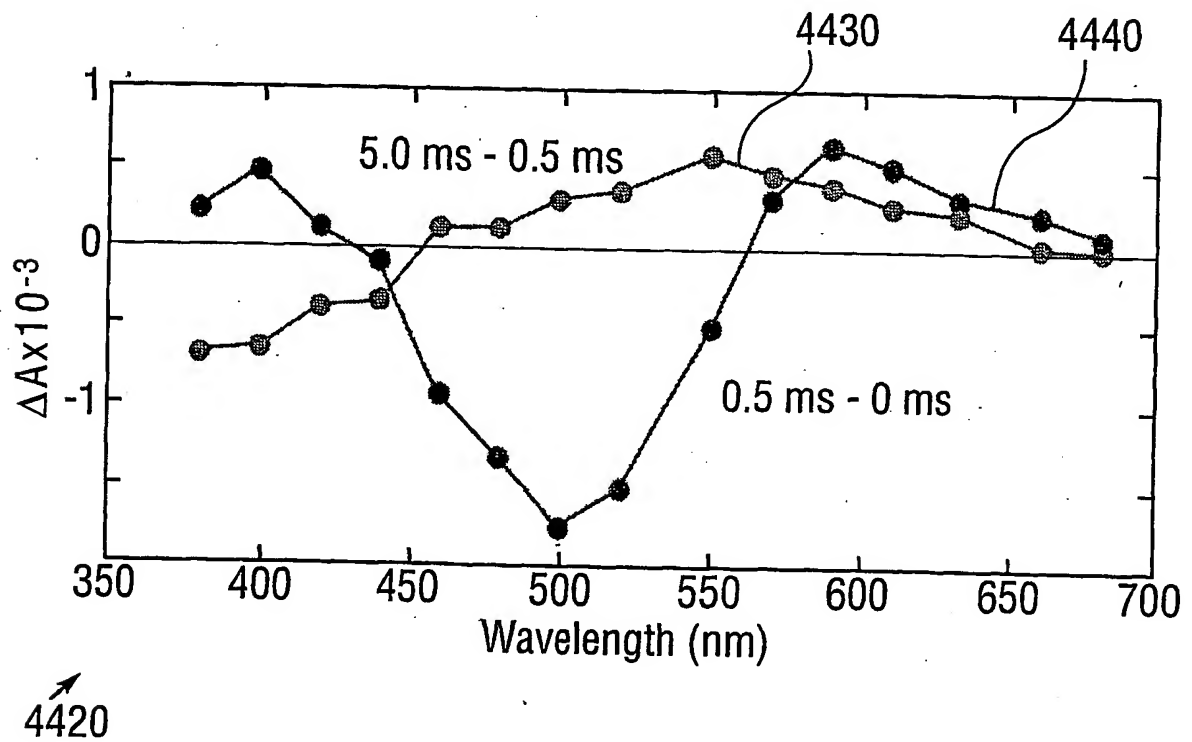
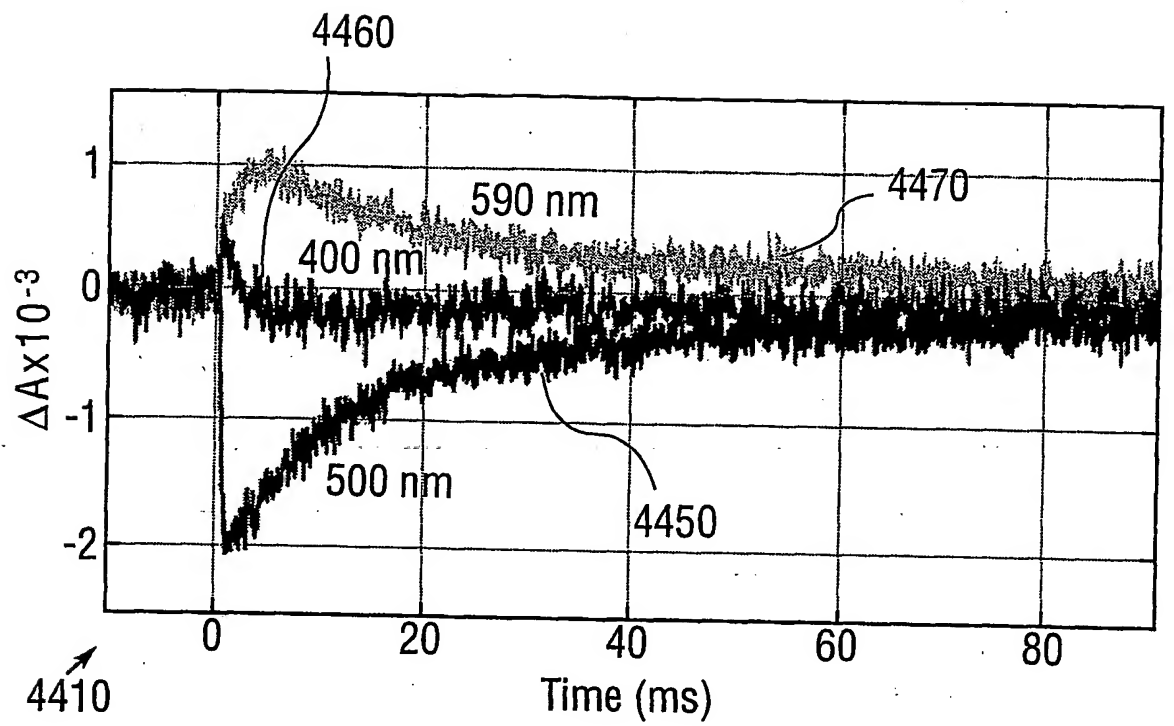


Fig. 44

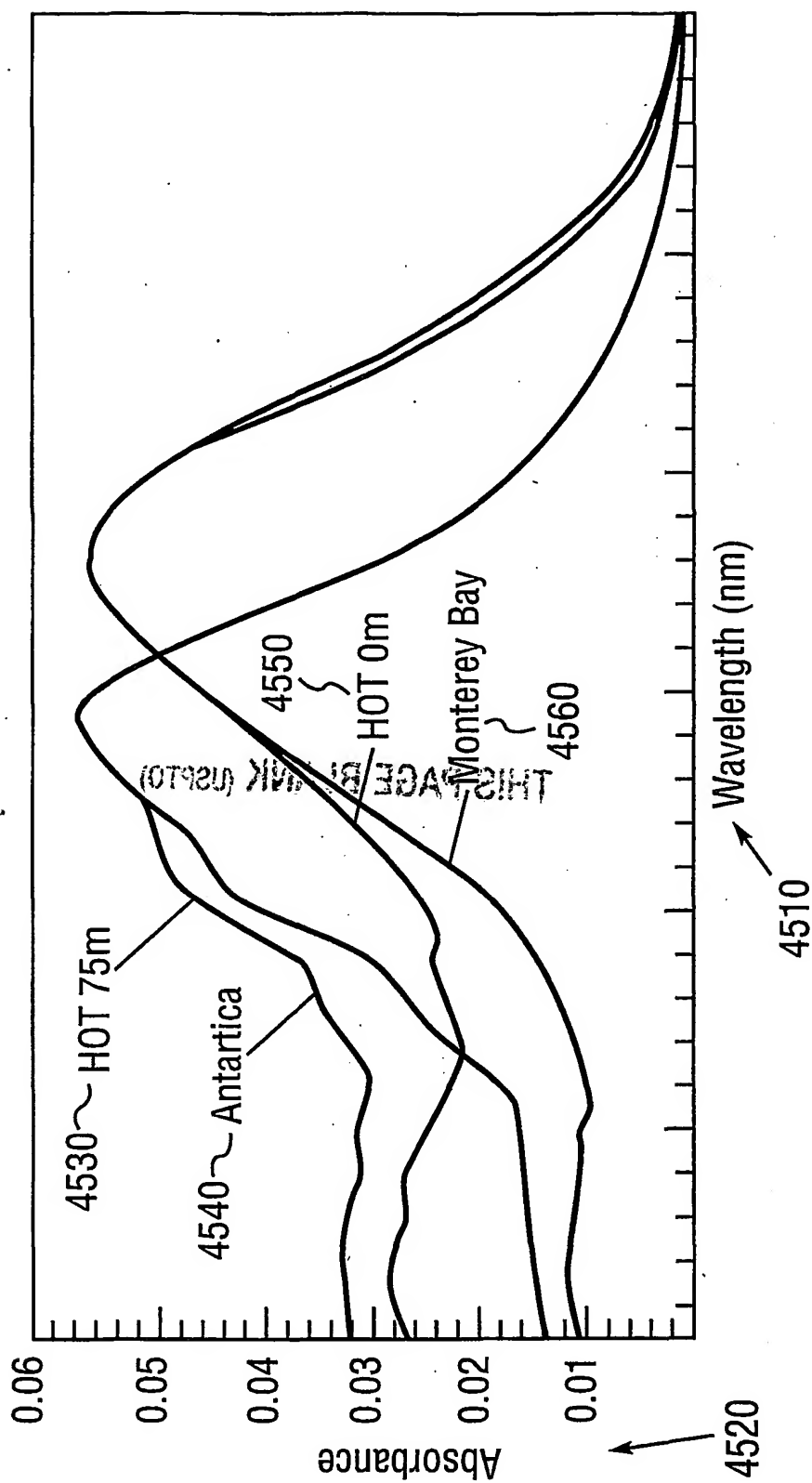


Fig. 45

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☒ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☐ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)